COMPUTATIONAL ANALYSES OF GENE EXPRESSION PROFILES OF OVARIAN AND PANCREATIC CANCER

A Dissertation
Presented to
The Academic Faculty

by

Loukia N. Lili

In Partial Fulfillment
of the Requirements for the Degree
of Philosophy in the
School of Biology

Georgia Institute of Technology
December 2013

COMPUTATIONAL ANALYSES OF GENE EXPRESSION PROFILES OF OVARIAN AND PANCREATIC CANCER

Approved by:

Dr. John F. McDonald, Advisor
School of Biology
Georgia Institute of Technology

Dr. Shafiq A. Khan
Department of Biological Sciences
Clark Atlanta University

Dr. King I. Jordan
School of Biology
Georgia Institute of Technology

Dr. Nathan J. Bowen
Department of Biological Sciences
Clark Atlanta University

Dr. George W. Daneker Jr., MD
Chief of Staff/Interim Chief of Surgery
Cancer Treatment Centers of America

Date Approved: [June 13th, 2013]

To my parents, Eleni Lili and Nikos Lilis for giving me a solid intellectual background

To my academic advisor Dr. John F. McDonald for guiding me through the adventurous paths of philosophy

To my life-time mentor, Dr. Thanasis Pheidas for supporting me in my academic and non-academic endeavors

To my partner, Richard D. Williams for being an excellent ally as a friend, as a mentor and as a triathlon coach in every step of my mental and physical efforts throughout my most crucial years of my graduate studies

ACKNOWLEDGEMENTS

First and foremost, I wish to thank my academic advisor Dr. John F. McDonald for everything he did to make me a strong, independent scientist. I also wish to thank Dr. Nathan Bowen for the valuable advice and guidance throughout my first year in the lab. I would like to thank Dr. DeEtte Walker and Dr. Lilya Matyunina for their professionalism in conducting all the laboratory experiments of my projects and for being part of my exciting academic family. I would also like to thank Dr. King I. Jordan for the priceless support on my first year in the Biology program and for any computational advice he provided for all of my projects on the following years. I also owe many thanks to Dr. Shafiq Khan for providing a different perspective on my projects, and Dr. George Daneker for enriching my knowledge in the clinical aspects of my research work. Last but not least, I wish to thank my PhD fellow students, Vinay Mittal, Chris Hill and Neda Jabbari for all the stimulating and philosophical discussions we had during my time in graduate school.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Cancer is a devastating disease for human society with thousands of deaths and estimated new cases every year around the globe. Intensive research efforts on understanding the disease progression and determining effective diagnostics and therapeutics have been employed for over one hundred years. Throughout this time, and in particular during the last two decades, computational-based methods have gained increasing importance in cancer biology research by providing significant advantages in the analysis and interpretation of high-throughput data at the molecular and genomic levels.

More specifically, after completion of the Human Genome Project in 2003, and with the Cancer Human Genome Project underway, high-throughput biological assays (e.g., microarray chips, next generation sequencing machines) have supplied researchers thousands of measurements per experimental sample. The massive amount of related data has oftentimes been challenging to interpret and translate, particularly in cancer biology and therapeutics. This thesis reports the results of three independent studies in which high-throughput gene expression is computationally analyzed to address longstanding issues in cancer biology. Two of the studies utilize data from ovarian cancer patients while the third involves data collected from pancreatic cancer patients.

In Chapter 1, I address the importance of personalized profiling in pancreatic cancer ; in Chapter 2 the role of cancer stroma in the progression of ovarian cancer and in Chapter 3 evidence for the role of epithelial-to-mesenchymal transition (EMT) in ovarian cancer metastasis.

More specifically, Chapter 1 emphasizes the power of personalized molecular profiling in unmasking unique gene expression signatures that correspond to each individual patient. These individual expression patterns (individual profiling), which may be overlooked by the traditional methods of gene signatures enriched in groups of afflicted individuals (group profiling), can provide valuable information for more successful targeted therapies. In order to address this issue in pancreatic cancer, comparisons of the most significantly differentially expressed genes and functional

pathways were performed between cancer and control patient samples as determined by group vs. personalized analyses. There was little to no overlap between genes/pathways identified by group analyses relative to those identified by personalized analyses. These results indicated that personalized and not group molecular profiling is the most appropriate approach for the identification of putative candidates for targeted gene therapy of pancreatic and perhaps other cancers with heterogeneous molecular etiology.

Chapter 2, also with strong implications on personalized molecular profiling, unveils the functional variability of the tumor microenvironment among ovarian cancer patients. The purpose of this study was to investigate the process of microenvironmental stroma activation in human ovarian cancer by molecular analysis of matched sets of cancer and surrounding stroma tissues from individual patients. Expression patterns of genes encoding signaling molecules and compatible receptors in the cancer stroma and cancer epithelia samples indicated the existence of two sub-groups of cancer stroma with different propensities to support tumor growth. These results demonstrated that functionally significant variability exists among ovarian cancer patients in the ability of the microenvironment to modulate cancer development.

Chapter 3 aims to uncover the molecular mechanisms that underlie the metastatic process with the hope that such knowledge may lead to more effective therapeutic treatments. For this purpose, pathological and molecular analyses were conducted in 14 matched sets of primary and metastatic samples from late staged ovarian cancer patients. Pathological examination revealed no morphological differences between any of the primary and metastatic samples. In contrast, gene expression analyses identified two distinct groups of patient samples. One group displayed essentially identical expression patterns to primary samples isolated from the same patients. The second group displayed expression patterns significantly different from primary samples isolated from the same patients. Predominant among the differentially expressed genes characterizing this second class of metastatic samples were genes previously associated with epithelial-to-mesenchymal transtion (EMT). These results supported a role of EMT in at least some ovarian cancer metastases and demonstrated that indistinguishable morphologies between primary and metastatic cancer samples is not sufficient evidence to negate the role of EMT in the metastatic process.

CHAPTER 1

EVIDENCE FOR THE IMPORTANCE OF PERSONALIZED MOLECULAR PROFILING IN PANCREATIC CANCER

1.1 Introduction

High-throughput molecular profiles (DNA and RNA sequencing, microarray gene expression analyses, *etc*.) are revolutionizing the way cancers are diagnosed [1-4], classified [5, 6] and treated [7-9]. One well-established approach to identify molecular variants (*e.g*., genetic, epigenetic, or gene expression pattern variants) that may be causally related to complex diseases such as cancer, is to identify variant patterns that are significantly enriched in groups of afflicted individuals relative to normal healthy controls. Examples of this approach are the various genome-wide association studies (GWAS) designed to identify disease-causing alleles [10, 11]. While the group approach can, by design, detect genetic or gene expression patterns that are in common among groups of afflicted individuals, genetic variants/molecular patterns that are unique to specific individuals, albeit of potential clinical significance, may go undetected using the group approach. This is likely to be especially true if there are multiple possible molecular paths to the same disease state as is believed to be the case for many, if not all, cancers [12].

In this study, we were interested in evaluating the impact of employing a group *vs.* a personalized approach in the analysis of gene expression profiles of a series of pancreatic cancer patients. We found that the most significant genes/molecular pathways identified among these patients, when analyzed as a group, were substantially different from the significant genes/molecular pathways identified when the analysis was

performed on an individual patient basis. Our results are consistent with earlier DNA sequence studies [13, 14] indicating that, on the molecular level, pancreatic cancer is a highly heterogeneous disease and, as a consequence, personalized gene expression profiling is critical to the acquisition of clinically significant information

1.2 Materials and Methods

1.2.1 Tissue collection and cell extraction

Patient tissues were collected at Northside Hospital (Atlanta, GA) under appropriate Institutional Review Board protocols. Following resection, the tumor tissues were examined by a pathologist and then placed in cryotubes and frozen in liquid nitrogen. Samples were transported on dry ice to Georgia Institute of Technology (Atlanta, GA), and stored at -80 ºC.

The tissue samples were examined microscopically and histology of ductal adenocarcinoma was verified by a pathologist. Following the examination and verification, tissue samples were embedded in cryomatrix (Shandon, Fisher Scientific, Pittsburg, PA, USA), and 7μm frozen sections were cut and attached to uncharged microscope slides. Immediately after dehydration and staining (HistoGene, LCM Frozen Section Staining Kit, Life Technologies, Carlsbad, CA, USA), slides were processed in an Autopix (Life Technologies) instrument for laser capture microdissection (LCM). For each of the four patients, three samples from their ductal epithelial tumor cells and three samples of their normal ductal epithelial cells were collected. All cells were isolated by LCM to ensure purity of samples. Approximately 30,000 cells were collected for each of the 24 total samples (12 cancer and 12 normal samples).

Table 1.1. Patient clinical information at the time of surgery and clinical outcome at the time of this study.

PATIENT	SEX/AGE	TUMOR STAGE	CLINICAL OUTCOME (months after surgery)
P1	Male/77	T3N0MX II	No Evidence of Disease (15)
P2	Male/69	T3NXMX II	Alive with Disease (16)
P3	Female/55	T3N1M0 II	No Evidence of Disease (8)
P4	Female/67	T3N1MX II	Distant Metastases (9)

1.2.2 RNA extraction and amplification

PicoPure RNA Isolation Kit (Life Technologies) protocols were followed for RNA extraction from the LCM cells on the Macro LCM caps in 30µL of extraction buffer. RNA quality was verified for all samples on the Bioanalyzer RNA Pico Chip (Agilent Technologies, Santa Clara, CA, USA). Total RNA from the above extractions was processed using Ovation® Pico WTA System (NuGEN) in conjunction with the Encore™ BiotinIL Module (NuGEN Technologies, San Carlos, CA, USA), to produce an amplified, biotin-labeled cDNA suitable for hybridizing to GeneChip Human Genome U133 Plus 2.0 Arrays (Affymetrix, Santa Clara, CA, USA) following manufacturer's recommendations.

1.2.3 Microarray data analysis

We generated 24 individual gene expression profiles from the three cancer and three normal biological replicate samples of the four patients. Affymetrix .CEL files were processed using the Affymetrix Expression Console (EC) Software Version 1.1 with the Robust Multi-Array Average (RMA) normalization method. The normalized expression values from all 24 samples were $\log_2$ transformed.

Group analysis: The initial data contained 54,675 probe set expression values of the Affymetrix Human Genome U133 Plus 2.0 chip. For the group analysis, the $\log_2$ transformed values were averaged across the 12 cancer and 12 normal samples. An unpaired t-test (p-value ≤ 0.005) was applied to identify those probe sets (350) that had significantly different expressions among all 12 cancer and all 12 normal samples. These 350 probe sets were employed in the group clustering analysis. Out of these 350 probe sets, the 287 unique, annotated genes were ranked by fold change (FC). The FC of each gene was calculated by subtracting the average normal value from the average cancer value. Pathway analyses were carried out using the web-based integrated software suite MetaCore of GeneGO (http://thomsonreuters.com/products_services/science/systems-biology/). Applying the default cutoff p-value ≤ 0.05, the 287 genes were found to be enriched for 22 pathways.

Individual patient analysis: For the individual patient analysis, the $\log_2$ transformed values were averaged across each individual's cancer and normal replicate samples. From each of the patient's initial 54,675 probe sets, an unpaired t-test (p-value $\leq$ 0.005) was applied to identify 188, 267, 435 and 291 probe sets that had significantly different expression between the cancer and normal replicate samples for each of the patients P1, P2, P3 and P4, respectively. As in the group analysis, these probe sets were employed in the individual clustering analyses (heat maps). Out of these, the 148, 221, 362 and 220 unique, annotated genes for P1, P2, P3 and P4, respectively, were ranked according to FC. The FC of each gene was calculated by subtracting the average normal value from the average cancer value for each individual. These genes also were used in the pathway analyses as described above (MetaCore GeneGO software suite). Applying

the default cutoff p-value ≤ 0.05, the genes were found to be significantly enriched for 15, 17, 25 and 30 pathways in P1, P2, P3 and P4, respectively. For the probe set clusterings (heat maps) in both the group and individual analyses, the $\log_2$ transformed values were normalized by Z-score statistics.

Analysis of data from the previously published study of Badea et al [15]: Seventy-eight Affymetrix .CEL files were downloaded from the GEO Omnibus database with accession number GDS4103. The files were processed using the Affymetrix Expression Console (EC) Software Version 1.1 with the Robust Multi-Array Average (RMA) normalization method, and the normalized expression values were $\log_2$ transformed, similarly to our sample analysis. All the 78 samples from 36 patients were used for the group analysis. For the individual analysis, the available two replicate cancer and two replicate normal samples from three patients were used (here-in referred to as patients P5, P6 and P7; 12 samples total). Both the group and the individual analyses were performed using the methods described above. Because technical replicates (multiple assays of the same biological patient sample) rather than biological replicates (assays of multiple biological samples from the same patient), as assayed in our study, were used in the Baeda *et al* study, the number of significantly differentiated genes at $p \leq 0.005$ was more than an order of magnitude greater than in our study. Thus, we employed an unpaired t-test with a more stringent cut off ($p \leq 0.00001$) than in our analysis in order to keep the number of significantly differentiated genes comparable to our study. Using this criterion, 17,658 significantly differentially expressed probe sets were detected, of which the 500 (330 annotated, unique genes) most significant were used for further analysis. For the

individual gene analysis, the same unpaired t-test with p-value ≤ 0.00001 identified 12, 37 and 22 significant probe sets (12, 29 and 20 annotated, unique genes) in patients P5, P6 and P7, respectively.

1.3 Results

1.3.1 Group profiling identifies genes and functional pathways previously implicated in pancreatic and other cancers

Genes: In the group profiling, all 12 cancer samples were compared against all 12 normal samples, and 350 probe sets (287 genes) were found to display significant differences in expression (p-value ≤ 0.005). The clustering of these 350 probe sets presented in Figure 1.1 demonstrates clear separation of the cancer and control samples. However, multiple samples taken from the same patient do not consistently cluster together indicating heterogeneity within both the cancer and control groups.

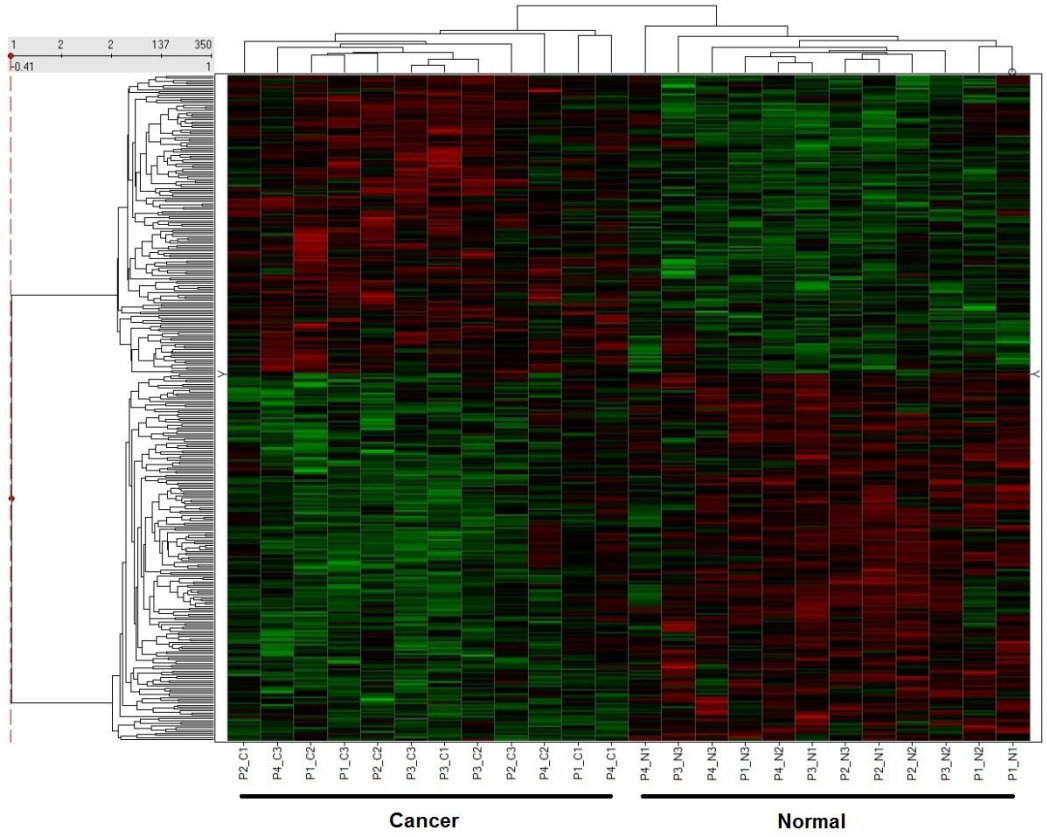


Figure 1.1. Hierarchical clustering of the 350 probe sets that display a significant (p-value $\leq$ 0.005) difference in gene expression among all Cancer and Normal samples. The heatmap was generated by z-score normalization of $\log_2$ expression values from the Affymetrix HG U133 Plus 2.0 chip. Patients are denoted as P1, P2, P3, and P4 with their associated cancer and normal biological replicate samples (*i.e.*, C1, C2, C3, and N1, N2, N3, respectively).

Table 1.2 presents the top 20 most significant differentially expressed genes (ten most significantly up-regulated and ten most significantly down-regulated) between the normal and cancer samples as ranked by FC (a complete listing of significantly differentiated genes is presented in Tables A.1.1 and A.1.2).

Table 1.2. The ten most significantly up-regulated and the ten most significantly down-regulated genes in cancer resulting from the group analysis of the 287 unique, annotated, significantly differentially expressed genes.

Probe Set ID	Gene Symbol	Fold Change	p-value
204351_at	S100P	2.514003	0.001932
242271_at	SLC26A9	2.178434	0.000706
219014_at	PLAC8	1.960672	0.000116
239196_at	ANKRD22	1.953161	0.004337
239609_s_at	LPCAT4	1.849687	0.00536
205769_at	SLC27A2	1.824208	0.001088
238021_s_at	CRNDE	1.768419	0.000247
58916_at	KCTD14	1.749122	0.000353
213611_at	AQP5	1.690754	0.000406
217109_at	MUC4	1.636676	0.000648
209277_at	TFPI2	-2.73483	0.001026
223761_at	FGF19	-2.67697	0.000627
204437_s_at	FOLR1	-2.52475	0.000341
1554690_a_at	TACC1	-2.15521	0.000136
214844_s_at	DOK5	-2.12474	0.004549
216598_s_at	CCL2	-2.1185	0.000506
223449_at	SEMA6A	-2.10883	0.001156
207392_x_at	UGT2B15	-1.99045	0.000764
204151_x_at	AKR1C1	-1.95521	0.00026
222901_s_at	KCNJ16	-1.93938	0.002889

Among the up-regulated genes in cancer was *S100P* (S100 calcium binding protein P), a member of the S100 protein family that has been associated with regulation of cell cycle progression and previously implicated in the etiology of prostate and pancreatic cancer [16, 17]. Also up-regulated was *ANKRD22* (ankyrin repeat domain 22), a gene previously found to be significantly up-regulated in the peripheral blood cells of

pancreatic cancer patients [18]. Other genes that were up-regulated in the cancer samples as determined in the group analysis are *MUC4* (mucin-4 cell surface associated), *CRNDE* (colorectal neoplasia differentially expressed), *and AQP5* (aquaporin 5). *MUC4* encodes a high-molecular weight glycoprotein that previously has been reported to be up regulated in pancreatic [19] and other cancers [20], and is believed to facilitate tumor growth and metastasis. *CRNDE* is a long non-coding RNA gene elevated in colorectal cancer [21].(*AQP5*) is a putative oncogene [22] that has been associated with increased proliferation and metastatic potential in breast cancer [23], lung cancer [24], non-small cell lung cancer [25], colorectal cancer [26], and chronic myelogenous leukemia [27]. Among the genes significantly down regulated in the cancer samples, *TFPI-2* (tissue factor pathway inhibitor 2) encodes a Kunitz-type serine proteinase inhibitor believed to regulate extracellular matrix digestion and remodeling [28]. Hyper-methylation of the *TFPI-2* promoter and consequent down-regulation in expression has been previously associated with the onset of pancreatic ductal adenocarcinomas [29] and other carcinomas [30]. Methylation of *TFPI-2* has also been recently proposed as a potential biomarker for the early detection of colorectal cancer [31]. Also significantly down regulated was *FGF19* (fibroblast growth factor 19), a member of the fibroblast growth factor (FGF) gene family that has been implicated in a variety of cancers [32]. Other genes found to be significantly down regulated in cancer tissues by the group analysis were *CCL2* (chemokine (C-C motif) ligand 2) *and TACC1* (transforming, acidic coiled-coil containing protein 1). *CCL2* is a chemokine that attracts mononuclear cells and possesses a dual role in cancer. In some instances, *CCL2* is involved in anti-tumor host activities, but in others it is secreted by cancer cells and it enhances cancer growth [33].

Loss of human *TACC* genes, *TACC1* and *TACC3*, underlies a significant portion of ovarian cancers [34].

Pathways: Functional analysis was carried out with the integrated software suite MetaCore of GeneGO (http://thomsonreuters.com/products_services/science/systems-biology/) incorporating the 287 differentially expressed genes. The analysis identified 22 significantly enriched functional pathways ($p \leq 0.05$, Table 1.3). More than half of the 22 pathways were associated with the immune response (12/22). Oncostatin M appeared in 4 out of the 12 immune response pathways. Oncostatin M is a member of a cytokine family that includes leukemia-inhibitory factor, granulocyte colony-stimulating factor, and interleukin 6, and it possesses the ability to inhibit the proliferation of cells in lines derived from several tumor types, including breast carcinoma, ovarian cancer, melanoma, glioma and lung carcinoma [35]. The two most significantly enriched pathways involve androstenedione and testosterone biosynthesis and metabolism (*i.e.*, androgen metabolism), both of which have been found significantly altered in pancreatic cancer [36]. Other immune response pathways from the group functional analysis were related to interleukins *IL-13, IL-17* and *IL-18*. *IL-13* was previously shown to play a pivotal role in the immunoregulatory pathway of NKT cells that suppress tumor immunosurveillance [37]. Although *IL-17* seems to have been previously associated with both tumor regression and tumor growth [38], the specific IL-17 immune response pathway enriched in our analysis contained the pro-tumorigenic gene, *CCL2*[33].

Table 1.3. The 22 significantly enriched pathways ($p \leq 0.05$) of the differentially expressed genes from the group analysis (287 genes).

GROUP ANALYSIS PATHWAYS	**p-Value**
Androstenedione and testosterone biosynthesis and metabolism p.2	0.000342
Androstenedione and testosterone biosynthesis and metabolism p.2/ Rodent version	0.000362
Immune response_Oncostatin M signaling via JAK-Stat in mouse cells	0.01474
Immune response_Oncostatin M signaling via JAK-Stat in human cells	0.01636
Regulation of lipid metabolism_FXR-dependent negative-feedback regulation of bile acids concentration	0.02527
Cell adhesion_Plasmin signaling	0.02849
Immune response_Oncostatin M signaling via MAPK in mouse cells	0.02849
Immune response_Oncostatin M signaling via MAPK in human cells	0.03009
HIV-1 signaling via CCR5 in macrophages and T lymphocytes	0.0317
Transport_ACM3 in salivary glands	0.0341
Immune response_IL-13 signaling via JAK-STAT	0.0357
Immune response_MIF-induced cell adhesion, migration and angiogenesis	0.03729
Development_GM-CSF signaling	0.04048
Immune response_Histamine signaling in dendritic cells	0.04048
Development_FGF-family signaling	0.04207
Immune response_CCL2 signaling	0.04366
Chemotaxis_CCL2-induced chemotaxis	0.04524
Immune response_TREM1 signaling pathway	0.04762
Triacylglycerol metabolism p.1	0.04762
Immune response_IL-17 signaling pathways	0.04841
Immune response_IL-18 signaling	0.04841
Immune response_CD40 signaling	0.05235

1.3.2 Personalized profiling identifies additional genes and functional pathways previously implicated in cancer

Genes: For the personalized profiles, the gene expression data for each individual patient were analyzed identically to the group profiling analyses. The number of

significantly differentially expressed probe sets between cancer and normal replicate samples of each patient (p-value ≤ 0.005) varied up to ~ 2-fold between patients (P1, 188 probe sets; P2, 267 probe sets; P3, 435 probe sets; P4, 291 probe sets). The clustering of these differentially expressed probe sets for each patient is presented as heat maps in Figure 1.2.

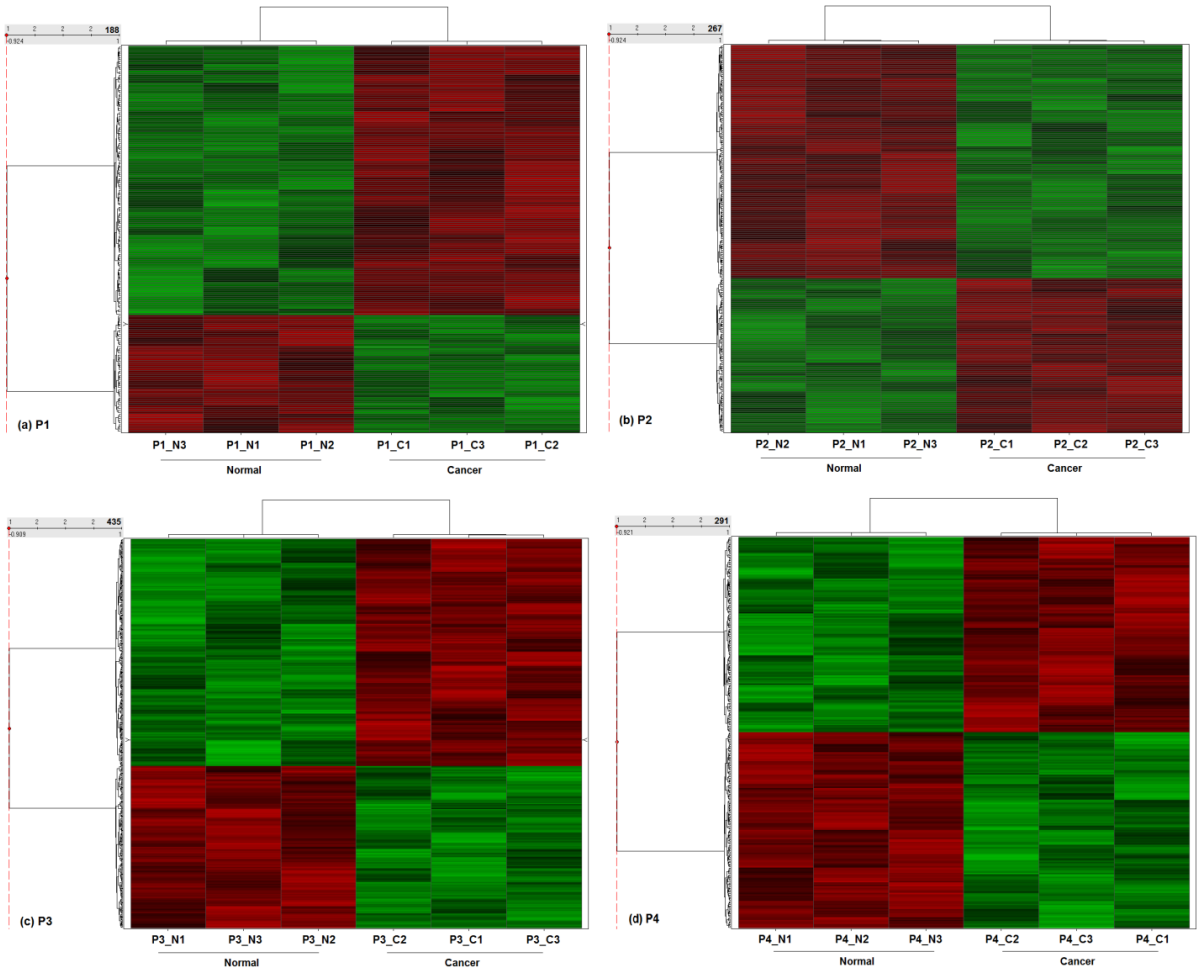


Figure 1.2. Supervised clustering of probe sets displaying a significant (p-value ≤ 0.005) difference in expression between normal and cancer samples for patients P1 (a), P2 (b), P3 (c), and P4 (d). The heatmap was generated by z-score normalization of $\log_2$ expression values from the Affymetrix HG U133 Plus 2.0 chip. Patients are denoted as P1, P2, P3, and P4 with their associated cancer and normal biological replicate samples (*i.e.*, C1, C2, C3, and N1, N2, N3, respectively).

A list of the 20 most significantly (p-value ≤ 0.005) differentially expressed genes ranked by fold change (10 most significantly up regulated and 10 most significantly

down regulated) between the normal and cancer samples for each individual patient are presented in Table 1.4 (a complete list of all significantly differentially expressed probe sets is presented in Tables A.1.1 and A.1.2).

Table 1.4. The ten most significantly (p-value ≤ 0.005) up-regulated and ten most down-regulated genes between normal and cancer samples for each patient P1, P2, P3 and P4 from the personalized profiling analysis (Figure 1.2).

Probe Set ID (Individual P1)	**Gene Symbol**	**Fold Change**	**p-value**	**Probe Set ID (Individual P2)**	**Gene Symbol**	**Fold Change**	**p-value**
205319_at	PSCA	3.732981	0.000613	220576_at	PGAP1	3.318862	0.001723
226517_at	BCAT1	2.959422	5.25E-05	217110_s_at	MUC4	3.103023	3.01E-05
1555294_a_at	ERC1	2.923769	0.000932	220133_at	ODAM	2.726326	0.003759
226325_at	ADSSL1	2.798377	0.001668	1567679_at	SNORA74A	2.636617	0.001267
228010_at	PPP2R2C	2.667305	0.001297	201926_s_at	CD55	2.627469	0.000945
52255_s_at	COL5A3	2.660518	0.001885	216504_s_at	SLC39A8	2.436613	0.002912
203877_at	MMP11	2.417413	0.002271	238022_at	CRNDE	2.275126	0.000994
205534_at	PCDH7	1.808023	0.002397	228962_at	PDE4D	2.273621	0.002435
207144_s_at	CITED1	1.525835	0.003498	212768_s_at	OLFM4	2.259513	0.000416
241368_at	PLIN5	1.471091	0.002259	205214_at	STK17B	2.178595	0.003404
1555236_a_at	PGC	-3.97182	0.000378	223761_at	FGF19	-4.87686	0.000303
219934_s_at	SULT1E1	-3.25992	0.002483	207016_s_at	ALDH1A2	-4.52882	0.004594
223509_at	CLDN2	-2.84658	0.000757	219106_s_at	KBTBD10	-4.32228	0.001409
226960_at	CXCL17	-2.76824	0.001549	209277_at	TFPI2	-4.17033	0.000529
201236_s_at	BTG2	-2.4767	0.005068	209993_at	ABCB1	-4.04187	0.000584
228912_at	VIL1	-2.40512	0.001675	204965_at	GC	-3.8697	0.003231
229254_at	MFSD4	-2.23402	0.00422	234673_at	HHLA2	-3.84065	3.20E-05
243296_at	NAMPT	-2.20159	0.000229	222257_s_at	ACE2	-3.54203	0.000581
1562625_at	FRYL	-2.08383	0.005056	205380_at	PDZK1	-3.37185	0.000588
225283_at	ARRDC4	-1.95739	0.003397	214397_at	MBD2	-3.23842	0.004437
Probe Set ID (Individual P3)	**Gene Symbol**	**Fold Change**	**p-value**	**Probe Set ID (Individual P4)**	**Gene Symbol**	**Fold Change**	**p-value**
204920_at	CPS1	4.650319	0.000671	215867_x_at	CA12	2.933468	0.00025

"Table continued"

239196_at	ANKRD22	4.62406	0.000438	208268_at	ADAM28	2.429627	0.004945
206291_at	NTS	4.048737	0.003113	235155_at	BDH2	1.938958	0.001707
220639_at	TM4SF20	3.977558	0.000243	229241_at	LDHD	1.735299	0.002328
218173_s_at	WHSC1L1	3.895598	1.04E-05	206242_at	TM4SF5	1.611106	0.000131
209806_at	HIST1H2BK	3.831179	0.003936	204602_at	DKK1	1.53691	0.003798
230252_at	LPAR5	3.741123	0.001988	224224_s_at	PDE11A	1.501315	0.001659
40020_at	CELSR3	3.695498	0.002774	1567079_at	CLN6	1.498309	0.00314
1557129_a_at	FAM111B	3.592115	0.000226	236129_at	GALNT5	1.434861	0.004932
1556357_s_at	ERICH1	3.544031	0.002244	219404_at	EPS8L3	1.428246	0.00397
214411_x_at	CTRB2	-5.7882	0.000797	219179_at	DACT1	-3.6374	0.001905
211766_s_at	PNLIPRP2	-5.52339	0.001025	213680_at	KRT6B	-3.3096	0.00037
207802_at	CRISP3	-5.49445	0.00027	206227_at	CILP	-2.94772	0.000918
205971_s_at	CTRB1///CTRB2	-5.42292	0.000386	204464_s_at	EDNRA	-2.86192	0.003475
205886_at	REG1B	-5.25456	0.000635	1560224_at	AHCTF1	-2.56906	0.000756
205509_at	CPB1	-5.24977	0.000438	226412_at	SFRS18	-2.51134	0.00462
209277_at	TFPI2	-4.04264	0.005253	225571_at	LIFR	-2.43013	0.005276
209616_s_at	CES1	-3.88049	0.003005	201108_s_at	THBS1	-2.42659	0.003524
207254_at	SLC15A1	-3.77115	0.000382	201838_s_at	SUPT7L	-2.42213	0.001748
211738_x_at	CELA3A	-3.54389	0.004419	1563321_s_at	MLLT10	-2.37822	0.004121

Patient 1 (P1). A number of genes previously associated with cancer onset/progression were among the top 20 most significantly differentially expressed in the P1 cancer sample. Among the most significantly up-regulated genes in P1 was *PSCA* (prostate stem cell antigen) a gene implicated in pancreatic cancer [39]. Also up regulated were *BCAT1* (branched chain amino-acid transaminase 1, cytosolic) and *MMP11* (matrix metallopeptidase 11 (stromelysin 3)). Over expression of *BCAT1* has been correlated with the clinical outcome of patients with breast cancer [40], colorectal cancer [41], neuroendocrine cancer [42] and melanoma [43]. Likewise, elevated *MMP11* expression

has been previously associated with aggressiveness of many cancer types [44, 45] and the MMP11 protein has been proposed as a target antigen for cancer immunotherapy [46].

Among the down-regulated genes in the P1 cancer sample were *NAMPT*, *VIL1*, *BTG2, CXCL17, CLDN2, SULT1E1* and *PGC*. *NAMPT* (nicotinamide phosphoribosyltransferase) is known to play an important role in the regulation of insulin secretion of pancreatic β cells and it has been characterized as an immunomodulatory gene that regulates cell viability, inflammation and cancer [47, 48]. Loss of *VIL1* (villin 1) is a feature of poorly differentiated colorectal cancers [49]. *BTG2* (BTG family, member 2), a p53-inducible anti-proliferative gene [50], has been reported to be absent in 65% of human breast tumors [51]. Interestingly, loss of *CXCL17* (chemokine (C-X-C motif) ligand 17) expression, as observed in P1, has been associated with progression from pancreatic adenoma to pancreatic adenocarcinoma [52]. The down-regulation of *CLDN2* (claudin 2) in P1 is consistent with the fact that most of the claudins display reduced expression in most cancers [53]. Lack of expression of the sulfotransferase family 1E, estrogen-preferring, member 1 (*SULT1E1*), a gene encoding an enzyme that catalyzes the sulfate conjugation of many hormones, neurotransmitters, drugs, and xenobiotic compounds, has been reported to be down regulated in prostate and breast cancer tissues and cell lines [54]. Lastly, reduced expression of *PGC* (progastricsin) has been associated with stomach cancer [55, 56].

Patient 2 (P2). Among the genes uniquely over-expressed in the P2 cancer sample was *ODAM* (ondotogenic ameloblast-associated), a gene that has recently been identified as a biomarker for breast cancer [57]. Over-expression of *CD55* (CD55 molecule, decay accelerating factor for complement (Cromer blood group)) has been associated with

prostate [58], breast [59], ovarian [60], and colorectal cancer [57]. Olfactomedin-4 (*OLFM4*) is a novel anti-apoptotic gene that has been reported to promote proliferation of pancreatic tumor cells [61]. Among the genes significantly down regulated in the cancer tissue isolated from patient 2, was *ALDH1A2* (aldehyde dehydrogenase 1 family, member A2). This gene has been identified as a tumor suppressor gene in prostate cancer [62]. *ABCB1* (ATP-binding cassette, sub-family B (MDR/TAP), member 1), also down regulated in P2 cancer, belongs to a family of genes that modulate the absorption, metabolism, and toxicity of pharmacological agents [63]. Down-regulation of *ACE2* (angiotensin I converting enzyme (peptidyl-dipeptidase A) 2) has been identified as a tumor suppressor in pancreatic cancer [64].

Patient 3 (P3). Among the uniquely over-expressed genes in P3 cancer was neurotensin (*NTS),* a gene encoding a secreted peptide widely distributed throughout the central nervous system, and recently shown to regulate growth of pancreatic cancer cells [65, 66]. Another up-regulated gene, *WHSC1L1* (Wolf-Hirschhorn syndrome candidate 1-like 1), is known to play a fundamental role in chromatin organization and is over-expressed in breast cancer [67, 68]. *CELSR3* (cadherin, EGF LAG seven-pass G-type receptor 3), also over-expressed in P3 cancer, is believed to be involved in contact-mediated communication (*i.e*., cell adhesions and ligand–receptor interactions) during cancer progression [69, 70]. Over-expression of *ERICH1* (glutamate-rich 1) has been associated with higher copy number in pancreatic cancer [71].

Among the uniquely down-regulated genes in P3 cancer, were *CTRB2* (chymotrypsinogen B2), *CPB1* (carboxypeptidase B1 (tissue) and *PNLIPRP2* (pancreatic lipase-related protein 2). Interestingly, over-expression of each of these genes has been

associated with poor-prognosis of pancreatic cancer patients [72]. The fact that these genes are down regulated in patient 3 may be related to this patient's relatively positive clinical outcome (Table 1.1). Also significantly down regulated were two genes, *CTRB1* (chymotrypsinogen B1) and *REG1B* (regenerating isle-derived 1 beta) that have been previously reported to be down regulated in pancreatic cancers [73].

Patient 4 (P4): The two most significantly over-expressed genes in P4, *CA12* and *ADAM28*, have been previously associated with tumor growth. The former, *CA12* (carbonic anydrase XII), has been hypothesized to maintain the acidic pH of tumors and therefore promote growth and invasion [74]. The latter, *ADAM28* (ADAM metallopeptidase domain 28), belongs to a novel gene family and has been reported to be over expressed in many malignant tumors promoting tumor growth and invasion [75, 76]. Among the genes down regulated in P4 cancer is *DACT1* and *KRT6B*. *DACT1* (dapper, antagonist of beta-catenin, homolog 1) plays an important role in stability of β catenin, and has been associated with colon cancer progression [77]. *KRT6B* (keratin 6B) encodes a basal-like cytokeratin that is aberrantly expressed in triple-negative breast cancer subtypes [78]. Of the remaining down-regulated genes, in P4 cancer is *EDNRA* (endothelin receptor type A), one of the endothelin receptors. There is recent evidence that down regulation of endothelin signaling results in cell survival alterations, cell invasiveness and carcinogenesis of various cancer types [79-81].

Pathways: As in the group analysis, functional pathway analysis was carried out on all significantly (p-value ≤ 0.005) differentially expressed, unique, annotated genes for each patient (P1, 148 genes; P2, 211 genes; P3, 351 genes; P4, 215 genes) to identify

functional pathways significantly overrepresented ($p \leq 0.05$) in the cancer samples isolated from each individual patient (Table 1.5).

Table 1.5. The significantly enriched pathways ($p \leq 0.05$) of the annotated, unique, differentially expressed genes in P1 (148 genes, 15 pathways), P2 (211 genes, 17 pathways), P3 (351 genes, 25 pathways) and P4 (215 genes, 30 pathways).

P1 Pathway Maps	**p-Value**	**P2 Pathway Maps**	**p-Value**
Immune response_NFAT in immune response	0.001482	Development_Notch Signaling Pathway	0.0004555
Immune response_CD28 signaling	0.001942	Transcription_Sin3 and NuRD in transcription regulation	0.003978
Cell adhesion_Tight junctions	0.01115	Cell cycle_Nucleocytoplasmic transport of CDK/Cyclins	0.006032
Immune response_TCR and CD28 co-stimulation in activation of NF-kB	0.01365	Development_Ligand-independent activation of ESR1 and ESR2	0.006414
Neurophysiological process_Glutamate regulation of Dopamine D1A receptor signaling	0.01708	Cytoskeleton remodeling_Integrin outside-in signaling	0.00813
Signal transduction_PKA signaling	0.02163	Signal transduction_PKA signaling	0.00908
Cell adhesion_ECM remodeling	0.02244	Development_Thrombopoetin signaling via JAK-STAT pathway	0.01466
Immune response_T cell receptor signaling pathway	0.02325	Cell adhesion_Endothelial cell contacts by non-junctional mechanisms	0.01732
Immune response_Immunological synapse formation	0.02839	Cell cycle_Regulation of G1/S transition (part 2)	0.02018
Glutathione metabolism	0.03299	Neurophysiological process_GABA-B receptor signaling at postsynaptic sides of synapses	0.02018
Cardiac Hypertrophy_NF-AT signaling in Cardiac Hypertrophy	0.03394	Immune response_Innate immune response to RNA viral infection	0.02322
Glutathione metabolism / Human version	0.03394	Chemotaxis_Leukocyte chemotaxis	0.0255
Glutathione metabolism / Rodent version	0.03886	LRRK2 in neurons in Parkinson's disease	0.03156
Chemotaxis_Leukocyte chemotaxis	0.04404	G-protein signaling_RhoA regulation pathway	0.03335
Development_Role of nicotinamide in G-CSF-induced granulopoiesis	0.05239	Cell adhesion_Plasmin signaling	0.03519
		Cell cycle_Regulation of G1/S transition (part 1)	0.04091
		Cell adhesion_Chemokines and adhesion	0.05265
P3 Pathway Maps	**p-Value**	**P4 Pathway Maps**	**p-Value**

"Table continued"

Cell cycle_Chromosome condensation in prometaphase	4.77E-06	Development_WNT signaling pathway. Part 2	0.00002776
Cell cycle_Spindle assembly and chromosome separation	0.000708	Apoptosis and survival_p53-dependent apoptosis	0.000972
Cell cycle_Transition and termination of DNA replication	0.004926	Mechanisms of CFTR activation by S-nitrosoglutathione (normal and CF)	0.00372
Proteolysis_Putative SUMO-1 pathway	0.005445	Cell cycle_Nucleocytoplasmic transport of CDK/Cyclins	0.003952
Cell cycle_Role of APC in cell cycle regulation	0.007195	Mechanism of Pioglitazone/ Metformin and Rosiglitazone/ Metformin cooperative action in Diabetes mellitus, Type 2	0.005167
Androstenedione and testosterone biosynthesis and metabolism p.2	0.009241	Cell cycle_Role of 14-3-3 proteins in cell cycle regulation	0.009688
Cell cycle_The metaphase checkpoint	0.009991	DNA damage_ATM / ATR regulation of G2 / M checkpoint	0.0134
Androstenedione and testosterone biosynthesis and metabolism p.2/ Rodent version	0.009991	Cell cycle_Transition and termination of DNA replication	0.01544
Immune response_IL-12-induced IFN-gamma production	0.009991	Development_Thrombospondin-1 signaling	0.01544
Transcription_Role of AP-1 in regulation of cellular metabolism	0.0116	DNA damage_Role of Brca1 and Brca2 in DNA repair	0.01762
Cell cycle_Nucleocytoplasmic transport of CDK/Cyclins	0.01272	Cell cycle_Role of APC in cell cycle regulation	0.01992
Transport_RAN regulation pathway	0.0207	LRRK2 in neurons in Parkinson's disease	0.02111
Cell cycle_Sister chromatid cohesion	0.03027	Cell cycle_Spindle assembly and chromosome separation	0.02111
Immune response_IL-12 signaling pathway	0.03289	Apoptosis and survival_Cytoplasmic/mitochondrial transport of proapoptotic proteins Bid, Bmf and Bim	0.02233
Glycolysis and gluconeogenesis p.3 / Human version	0.03559	Estradiol metabolism	0.02358
Glycolysis and gluconeogenesis p.3	0.03559	Estrone metabolism	0.02358
Cell cycle_Initiation of mitosis	0.03837	Estradiol metabolism / Human version	0.02486
Immune response_IL-23 signaling pathway	0.03837	Estrone metabolism / Human version	0.02486
DNA damage_ATM / ATR regulation of G2 / M checkpoint	0.04124	Estradiol metabolism / Rodent version	0.02617
Glutathione metabolism / Human version	0.04722	Cell cycle_Regulation of G1/S transition (part 1)	0.0275
Glycolysis and gluconeogenesis (short map)	0.04904	Cell adhesion_Chemokines and adhesion	0.0306

"Table continued"

Cell cycle_Role of SCF complex in cell cycle regulation	0.0503	Apoptosis and survival_BAD phosphorylation	0.03311
Apoptosis and survival_p53-dependent apoptosis	0.0503	Signal transduction_AKT signaling	0.03457
Apoptosis and survival_Granzyme A signaling	0.05346	Apoptosis and survival_FAS signaling cascades	0.03606
Cytoskeleton remodeling_RalA regulation pathway	0.05346	Development_Adiponectin signaling	0.03758
		Cytoskeleton remodeling_TGF, WNT and cytoskeletal remodeling	0.03985
		NAC-AsPC-1	0.04682
		NAC-AsPC-1	0.04682
		Untitled	0.04682
		Some pathways of EMT in cancer cells	0.04716

Patient 1 (P1). Five of the 15 most significantly enriched pathways in P1 are associated with the immune response. More specifically, *NFAT* (nuclear factor of activated T cells) is a major transcriptional regulator in T cells and recently identified as a potent immuno-regulator in cancer development and as a potential target for therapeutic manipulation of the immune response in cancer patients [82]. P1 also showed enrichment for the TCR and CD28 signaling pathways. Glutathione metabolism was also identified as a significantly enriched pathway in P1. Glutathione is known to affect the efficacy of anti-neoplastic interventions mainly through nucleophilic thioether formation or oxidation-reduction reactions [83]. The prevalence of enriched immune response and glutathione metabolism pathways may help account, thus far, for the favorable outcome in patient P1 (Table 1.1).

Patient 2 (P2). Patient P2 displayed pathways that have been implicated strongly in cancer development and invasion. Notch signaling participates in many developmental processes regulating cell differentiation, proliferation, apoptosis, adhesion, epithelial-to-

mesenchymal transition, migration, angiogenesis, and can act either as an oncogene or tumor suppressor in a highly context-dependent manner [84]. Cell-cycle disruption is a typical feature of cancer cells and results in DNA damage [85]. Cytoskeleton remodeling is required for cancer cell invasion and metastasis, apparent in most cancers [86]. Cell adhesion determines the polarity of cells and maintains the cell architecture in tissues. Cell adhesiveness is generally reduced in cancer to allow for invasiveness, extra cellular matrix decomposition and metastasis [87].

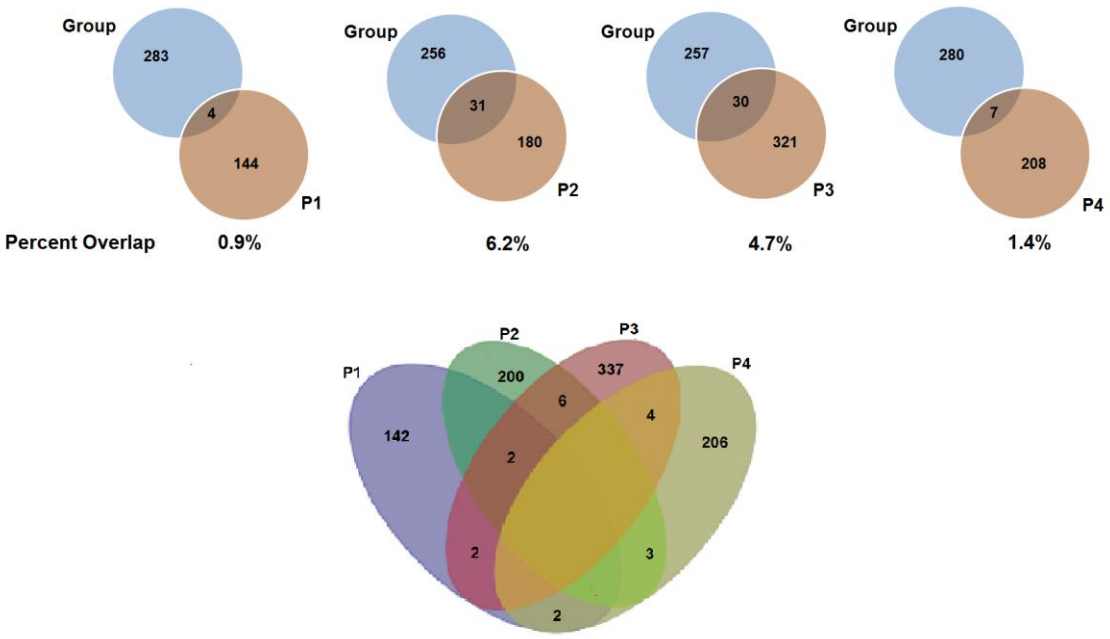
Patient 3 (P3). Genes in patient P3 were enriched predominantly for cell cycle regulatory pathways (nine out of a total 25 pathways). This is typical for cancer cells at an advanced stage as with patient P2. Like patient P1, P3 showed enrichment of interleukin-mediated immune responses, and the glutathione metabolism pathway. Interleukin-12 (*IL-12*) is a powerful coordinator of the innate and adaptive immune responses and has been shown to have promising anti-tumor effects in murine tumor models [88]. *IL-12* is currently being investigated as a potential therapeutic agent against cancer [89].

Patient 4 (P4). The most significantly enriched pathway in P4 was the WNT signaling pathway. The canonical WNT/□-catenin pathway has emerged as a critical regulator in stem cells and has also been associated with cancer in many tissues [90]. For P4, this particular WNT pathway, involved the frizzled family receptor 7 *(FZD7),* which was up regulated. Up regulation of *FZD7* has been reported in gastric and colorectal cancers [91, 92]. P4 also showed enrichment of apoptotic and survival pathways. In the p53 dependent apoptosis pathway, the *BCL2L11* gene (BCL2-like 11-apoptosis facilitator), responsible for cytoplasmic transport of proapoptotic proteins BID, BMF and

BIM, is down regulated. On the other hand, *CDK1* (cyclin-dependent kinase 1) that promotes phosphorylation of the proapoptotic *BAD* (BCL-2-associated agonist of cell death) was up regulated in the BAD phosphorylation pathway. This is evidence for de-regulation of the apoptosis and survival pathways in patient P4.

1.3.3 Significant genes and pathways in the personalized analyses display little to no overlap among individual patients or with those identified in the group analysis

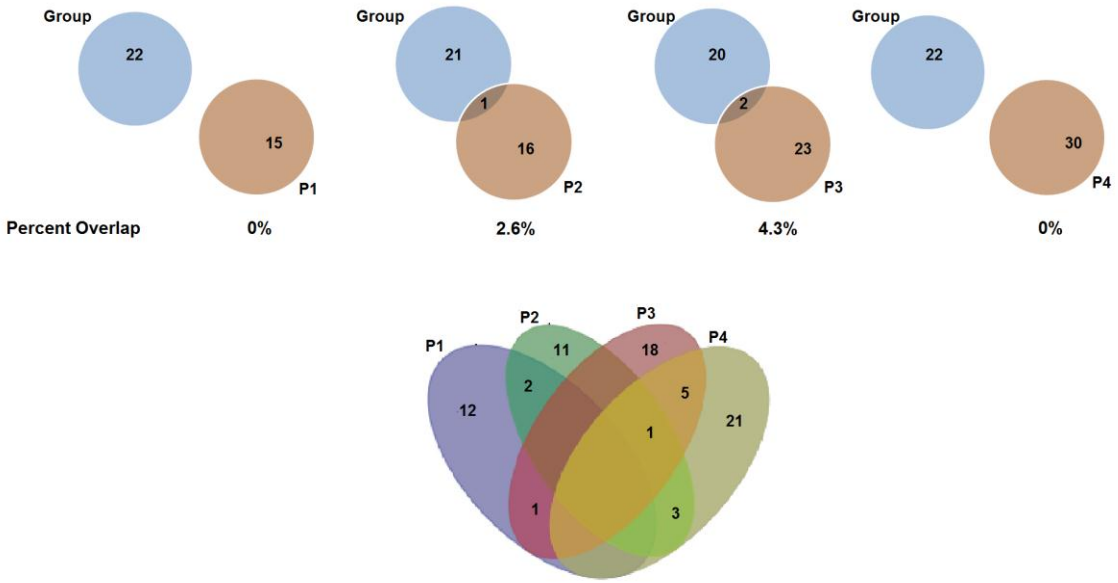
As shown above, both the group and the personalized analyses identified genes and pathways previously implicated in the onset/progression of pancreatic and a broad spectrum of other cancers. We were next interested in determining the degree of overlap among those genes and pathways identified as significant in each of the individual patient analyses and in the group analysis. Interestingly, we found that the degree of overlap is remarkably low. As shown in Figure 1.3 (see also Tables A.1.1 and A.1.2), less than 6.5% (average 3.3%) of the genes identified as significantly differentially expressed between normal and cancer cells isolated from individual patients (personalized profiles) overlap with genes identified as significantly differentially expressed across the combined patient samples (group analysis). Likewise, there is remarkably little overlap among the individual patients. For example, of the combined number of annotated genes identified as significantly differentially expressed in samples P1 and P2 (148 + 211 = 359), there was less than 1% ($2/359 \approx 0.006$) overlap. Even between P2 and P3, samples that share the largest number of overlapping genes (8 genes), the degree of overlap is only slightly more than 1% ($8/(211+351) \approx 0.014$).



	Group	P1	P2	P3	P4
Group	-	4	31	30	7
P1		-	2	4	2
P2			-	8	3
P3				-	4
P4					-

Figure 1.3. Venn diagrams and table showing the unique, annotated genes identified as significantly differentially expressed in the group analysis and in the personalized analysis(es) of at least one patient.

Comparison of the most significantly over-represented pathways identified in the personalized and group analyses resulted in similar results to the gene analyses, *i.e.*, there is relatively little overlap between pathways identified as over-represented in the group analysis vs. the personalized analyses. Furthermore, there is remarkably little overlap in over-represented pathways among individual patients based on the personalized profiles (Figure 1.4, Table A.1.3).



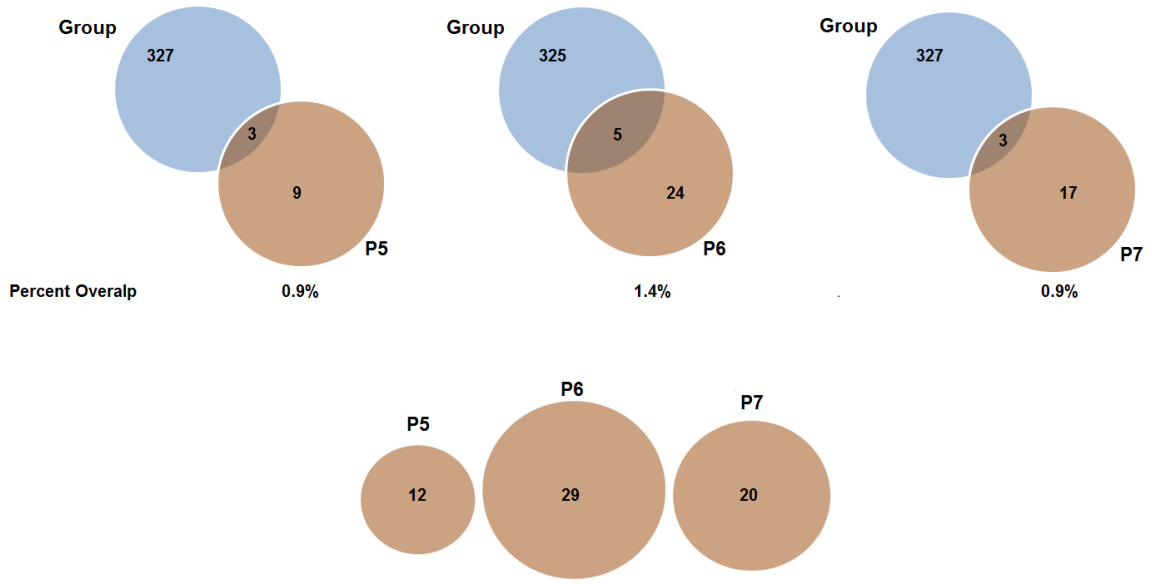
	Group	P1	P2	P3	P4
Group	-	0	1	2	0
P1		-	2	1	0
P2			-	1	4
P3				-	6
P4					-

Figure 1.4. Venn diagrams and table showing pathways identified as significantly enriched in the group analysis and in the personalized analysis(es) of at least one patient.

As shown in Figure 1.4 (see also Table A.1.3), less than 5% (average 1.7%) of the pathways identified as significantly over-represented in individual patients (personalized profiles) overlap with pathways of genes identified as significantly differentially expressed across the combined patient samples (group analysis). In fact, pathways identified as over-represented in two of the patient samples (P1 and P4), had no overlap with those identified in the group analysis. Additionally, there is relatively little overlap

among the individual patients. For example, of the pathways identified as significantly over-represented in samples P1 and P2 (15 + 17 = 32), there was only 6.3% (2/32 ≈ 0.063) overlap. Even between P3 and P4, samples that share the largest number of significantly over-represented pathways (6 pathways), the degree of overlap is less than 11% (6/(25 + 30) ≈ 0.109).

The results of the above studies indicate that genes and pathways identified as being most significantly different between normal and cancer samples as determined by the group analysis display little or no overlap with those identified as significant by individual personalized analyses. Likewise, we found little or no overlap in genes and pathways identified as being most significantly different among individual patient samples (personalized analyses). To address the possibility that these findings may simply be an artifact of the relatively small number of patients examined in our study, we conducted a similar analysis using data from a previously published microarray gene expression analysis of control and cancer tissue samples isolated from 36 patients [15]. In this earlier study, replicate assays were carried out on three patients allowing us to compare the most significantly differentiated genes as determined by a group analysis (36 patient samples) vs. the significantly differentiated genes determined in personalized analyses of three patients. Consistent with our previous findings, the results demonstrate remarkably little overlap between genes identified as significant in the group vs. personalized analyses (Figure 1.5, Table A.1.4).



	Group	P5	P6	P7
Group	-	3	5	3
P5		-	0	0
P6			-	0
P7				-

Figure 1.5. Venn diagrams and table showing the genes identified as significantly differentially expressed in the group analysis and in the personalized analysis(es) of at least one patient (data from Badea *et. al.*).

As shown in Figure 1.5 (see also Table A.1.4), less than 2% (average 1.07%) of the genes identified as significantly differentially expressed between normal and cancer cells ($p \leq 0.00001$) isolated from individual patients (personalized profiles) overlap with genes identified as most significantly differentially expressed (top 500 of 17,658 genes significantly differentially expressed, $p \leq 0.00001$) across the combined patient samples. There was no overlap among patients in significantly differentiated genes.

1.4 Discussion

Molecular profiling is revolutionizing the way we view and treat cancer. Rather than the traditional tissue-of-origin approach to the classification and treatment of the disease, molecular profiling is providing gene-based diagnostics and therapeutics as a realistic alternative. The identification of key genes/pathways associated with various types of cancer is the foundation for both molecular diagnostics and therapeutics. The group approach to the identification of key genes/pathways involves combining the molecular profiles of collections of samples from diseased patients in order to identify shared variant profiles that are distinct from those associated with non-diseased controls (*e.g.*, [93]) . While this can be a productive approach for the detection of biomarkers and potential therapeutic targets for diseases caused by one or a few genes, for diseases caused by aberrations in a variety of alternative genes/pathways, the group approach may be less effective [94].

Initial sequence analyses of tumor samples isolated from a large number of cancer patients suggest that, from the molecular perspective, pancreatic cancer may be a highly heterogeneous disease [13]. In the present study, we were interested in further examining this question by comparing the most significantly differentially expressed genes/pathways between pancreatic cancer and control samples as determined by group vs. personalized analyses of the same samples. Toward this end, we employed LCM to collect three distinct sets (biological replicates) of normal and cancer cells from tissue samples obtained from four pancreatic patients. In addition, we re-analyzed data from a previous gene expression analysis of 36 pancreatic patients [15] and compared the most significantly differentiated genes/pathways as determined by the group analysis relative

to the most significantly differentiated genes/pathways as determined by personalized analyses of three patients for which replicate microarray assays were performed.

Our results consistently demonstrated little to no overlap between genes/pathways identified in the group analyses relative to those identified in the personalized analyses. For example, consistent with earlier reports [95], our group analysis identified *MUC4* as one of the most significantly differentiated expressed genes between the normal and pancreatic cancer samples (Table 1.2). Indeed, MUC4 has recently been proposed as a prime candidate for targeted drug therapy in pancreatic cancer [20]. In our personalized analyses, however, MUC4 was identified as significantly over-expressed in only 1 of the 7 patients examined suggesting that MUC4 therapy would likely not be effective for the majority of the patients examined in our study. Conversely, many of the genes identified as being significantly differentially expressed in individual patients (personalized profiles) were not identified as significant in the group analysis. For example, the most significantly differentially expressed gene in the cancer samples isolated from patient 1 is *PSCA* (prostate stem cell antigen). Interestingly, a monoclonal antibody against PSCA is currently being tested in clinical trials for both prostate and pancreatic cancer [96, 97]. Thus, while PSCA targeted therapy might well be expected to be effective for patient 1, it was not identified as being significantly over-expressed in the group analysis nor in the personalized analyses of any of the other patients examined. Similarly, *ADAM* (a disintegrin and metalloprotease), a gene reported to be over-expressed in a number of human cancers [75] and identified as a potential candidate for targeted gene therapy [98], was among the most significantly over-expressed genes in patient 4 but was not identified

as being significantly over-expressed in the group analysis nor in the personalized analyses of any of the other patients examined.

Collectively, our results are consistent with earlier findings indicating that, on the molecular level, pancreatic cancer is a highly heterogeneous disease [13]. While targeted gene therapy holds great promise in the treatment of pancreatic and other cancers, a crucial step in the process is the accurate identification of appropriate candidate genes for targeted therapy. Our findings indicate that personalized and not group molecular profiling is the most appropriate approach for the identification of putative candidates for effective targeted gene therapy for pancreatic and perhaps other cancers with heterogeneous molecular etiology.

1.5 References

1. Idris, S.F., et al., *The role of high-throughput technologies in clinical cancer genomics.* Expert review of molecular diagnostics, 2013. **13**(2): p. 167-81

2. Lakhani, S.R. and A. Ashworth, *Microarray and histopathological analysis of tumours: the future and the past?* Nature reviews. Cancer, 2001. **1**(2): p. 151-7.

3. ten Bosch, J.R. and W.W. Grody, *Keeping up with the next generation: massively parallel sequencing in clinical diagnostics.* The Journal of molecular diagnostics : JMD, 2008. **10**(6): p. 484-92.

4. Su, Z., et al., *Next-generation sequencing and its applications in molecular diagnostics.* Expert review of molecular diagnostics, 2011. **11**(3): p. 333-43.

5. Sotiriou, C., et al., *Breast cancer classification and prognosis based on gene expression profiles from a population-based study.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(18): p. 10393-8.

6. Chuang, H.Y., et al., *Network-based classification of breast cancer metastasis.* Molecular systems biology, 2007. **3**: p. 140.

7. Meyerson, M., S. Gabriel, and G. Getz, *Advances in understanding cancer genomes through second-generation sequencing.* Nature reviews. Genetics, 2010. **11**(10): p. 685-96.

8. van't Veer, L.J. and R. Bernards, *Enabling personalized cancer medicine through analysis of gene-expression patterns.* Nature, 2008. **452**(7187): p. 564-70.

9. Wulfkuhle, J., et al., *Genomic and proteomic technologies for individualisation and improvement of cancer treatment.* European journal of cancer, 2004. **40**(17): p. 2623-32.

10. Altshuler, D., M.J. Daly, and E.S. Lander, *Genetic mapping in human disease.* Science, 2008. **322**(5903): p. 881-8.

11. McClellan, J. and M.C. King, *Genetic heterogeneity in human disease.* Cell, 2010. **141**(2): p. 210-7.

12. Vogelstein, B. and K.W. Kinzler, *Cancer genes and the pathways they control.* Nature medicine, 2004. **10**(8): p. 789-99.

13. Jones, S., et al., *Core signaling pathways in human pancreatic cancers revealed by global genomic analyses.* Science, 2008. **321**(5897): p. 1801-6.

14. Liang, W.S., et al., *Genome-wide characterization of pancreatic adenocarcinoma patients using next generation sequencing.* PloS one, 2012. **7**(10): p. e43192.

15. Badea, L., et al., *Combined gene expression analysis of whole-tissue and microdissected pancreatic ductal adenocarcinoma identifies genes specifically overexpressed in tumor epithelia.* Hepato-gastroenterology, 2008. **55**(88): p. 2016-27.

16. Arumugam, T., et al., *S100P promotes pancreatic cancer growth, survival, and invasion.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2005. **11**(15): p. 5356-64.

17. Basu, G.D., et al., *Functional evidence implicating S100P in prostate cancer progression.* International journal of cancer. Journal international du cancer, 2008. **123**(2): p. 330-9.

18. Baine, M.J., et al., *Transcriptional profiling of peripheral blood mononuclear cells in pancreatic cancer patients identifies novel genes with potential diagnostic utility.* PloS one, 2011. **6**(2): p. e17014.

19. Chaturvedi, P., et al., *MUC4 mucin interacts with and stabilizes the HER2 oncoprotein in human pancreatic cancer cells.* Cancer research, 2008. **68**(7): p. 2065-70.

20. Singh, A.P., P. Chaturvedi, and S.K. Batra, *Emerging roles of MUC4 in cancer: a novel target for diagnosis and therapy.* Cancer research, 2007. **67**(2): p. 433-6.

21. Graham, L.D., et al., *Colorectal Neoplasia Differentially Expressed (CRNDE), a Novel Gene with Elevated Expression in Colorectal Adenomas and Adenocarcinomas.* Genes & cancer, 2011. **2**(8): p. 829-40.

22. Woo, J., et al., *Overexpression of AQP5, a putative oncogene, promotes cell growth and transformation.* Cancer letters, 2008. **264**(1): p. 54-62.

23. Jung, H.J., et al., *Aquaporin-5: a marker protein for proliferation and migration of human breast cancer cells.* PloS one, 2011. **6**(12): p. e28492.

24. Zhang, Z., et al., *Expression of aquaporin 5 increases proliferation and metastasis potential of lung cancer.* The Journal of pathology, 2010. **221**(2): p. 210-20.

25. Chae, Y.K., et al., *Expression of aquaporin 5 (AQP5) promotes tumor invasion in human non small cell lung cancer.* PloS one, 2008. **3**(5): p. e2162.

26. Kang, S.K., et al., *Role of human aquaporin 5 in colorectal carcinogenesis.* The American journal of pathology, 2008. **173**(2): p. 518-25.

27. Chae, Y.K., et al., *Human AQP5 plays a role in the progression of chronic myelogenous leukemia (CML).* PloS one, 2008. **3**(7): p. e2594.

28. Chand, H.S., D.C. Foster, and W. Kisiel, *Structure, function and biology of tissue factor pathway inhibitor-2.* Thrombosis and haemostasis, 2005. **94**(6): p. 1122-30.

29. Sato, N., et al., *Epigenetic inactivation of TFPI-2 as a common mechanism associated with growth and invasion of pancreatic ductal adenocarcinoma.* Oncogene, 2005. **24**(5): p. 850-8.

30. Sierko, E., M.Z. Wojtukiewicz, and W. Kisiel, *The role of tissue factor pathway inhibitor-2 in cancer biology.* Seminars in thrombosis and hemostasis, 2007. **33**(7): p. 653-9.

31. Glockner, S.C., et al., *Methylation of TFPI2 in stool DNA: a potential novel biomarker for the detection of colorectal cancer.* Cancer research, 2009. **69**(11): p. 4691-9.

32. Turner, N. and R. Grose, *Fibroblast growth factor signalling: from development to cancer.* Nature reviews. Cancer, 2010. **10**(2): p. 116-29.

33. Conti, I. and B.J. Rollins, *CCL2 (monocyte chemoattractant protein-1) and cancer.* Seminars in cancer biology, 2004. **14**(3): p. 149-54.

34. Lauffart, B., et al., *Aberrations of TACC1 and TACC3 are associated with ovarian cancer.* BMC women's health, 2005. **5**: p. 8.

35. Grant, S.L. and C.G. Begley, *The oncostatin M signalling pathway: reversing the neoplastic phenotype?* Molecular medicine today, 1999. **5**(9): p. 406-12.

36. Fernandez-del Castillo, C., et al., *Pancreatic cancer and androgen metabolism: high androstenedione and low testosterone serum levels.* Pancreas, 1990. **5**(5): p. 515-8.

37. Terabe, M., J.M. Park, and J.A. Berzofsky, *Role of IL-13 in regulation of anti-tumor immunity and tumor growth.* Cancer immunology, immunotherapy : CII, 2004. **53**(2): p. 79-85.

38. Murugaiyan, G. and B. Saha, *Protumor vs antitumor functions of IL-17.* Journal of immunology, 2009. **183**(7): p. 4169-75.

39. Argani, P., et al., *Discovery of new markers of cancer through serial analysis of gene expression: prostate stem cell antigen is overexpressed in pancreatic adenocarcinoma.* Cancer research, 2001. **61**(11): p. 4320-4.

40. Karlsson, E., et al., *Gene expression variation to predict 10-year survival in lymph-node-negative breast cancer.* BMC cancer, 2008. **8**: p. 254.

41. Yoshikawa, R., et al., *ECA39 is a novel distant metastasis-related biomarker in colorectal cancer.* World journal of gastroenterology : WJG, 2006. **12**(36): p. 5884-9.

42. He, P., et al., *Identification of carboxypeptidase E and gamma-glutamyl hydrolase as biomarkers for pulmonary neuroendocrine tumors by cDNA microarray.* Human pathology, 2004. **35**(10): p. 1196-209.

43. van Houdt, I.S., et al., *Expression of the apoptosis inhibitor protease inhibitor 9 predicts clinical outcome in vaccinated patients with stage III and IV melanoma.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2005. **11**(17): p. 6400-7.

44. Egeblad, M. and Z. Werb, *New functions for the matrix metalloproteinases in cancer progression.* Nature reviews. Cancer, 2002. **2**(3): p. 161-74.

45. Foda, H.D. and S. Zucker, *Matrix metalloproteinases in cancer invasion, metastasis and angiogenesis.* Drug discovery today, 2001. **6**(9): p. 478-482.

46. Peruzzi, D., et al., *MMP11: a novel target antigen for cancer immunotherapy.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2009. **15**(12): p. 4104-13.

47. Garten, A., et al., *Nampt: linking NAD biology, metabolism and cancer.* Trends in endocrinology and metabolism: TEM, 2009. **20**(3): p. 130-8.

48. Galli, M., et al., *The nicotinamide phosphoribosyltransferase: a molecular link between metabolism, inflammation, and cancer.* Cancer research, 2010. **70**(1): p. 8-11.

49. Arango, D., et al., *Villin expression is frequently lost in poorly differentiated colon cancer.* The American journal of pathology, 2012. **180**(4): p. 1509-21.

50. Rouault, J.P., et al., *Identification of BTG2, an antiproliferative p53-dependent component of the DNA damage cellular response pathway.* Nature genetics, 1996. **14**(4): p. 482-486.

51. Kawakubo, H., et al., *Expression of the NF-kappaB-responsive gene BTG2 is aberrantly regulated in breast cancer.* Oncogene, 2004. **23**(50): p. 8310-9.

52. Hiraoka, N., et al., *CXCL17 and ICAM2 are associated with a potential anti-tumor immune response in early intraepithelial stages of human pancreatic carcinogenesis.* Gastroenterology, 2011. **140**(1): p. 310-21.

53. Hewitt, K.J., R. Agarwal, and P.J. Morin, *The claudin gene family: expression in normal and neoplastic tissues.* BMC cancer, 2006. **6**: p. 186.

54. Falany, C.N., et al., *Expression and subcellular localization of human sulfotransferases (SULTs) in normal and cancerous prostate and breast tissues.* Proceedings of the American Association for Cancer Research, 2004. **2004**(1): p. 1020.

55. Kim, B., et al., *Expression profiling and subtype-specific expression of stomach cancer.* Cancer research, 2003. **63**(23): p. 8248-55.

56. Hasegawa, S., et al., *Genome-wide analysis of gene expression in intestinal-type gastric cancers using a complementary DNA microarray representing 23,040 genes.* Cancer research, 2002. **62**(23): p. 7012-7.

57. Siddiqui, S., et al., *Odontogenic ameloblast associated protein as a novel biomarker for human breast cancer.* The American surgeon, 2009. **75**(9): p. 769-75; discussion 775.

58. Loberg, R.D., et al., *Inhibition of decay-accelerating factor (CD55) attenuates prostate cancer growth and survival in vivo.* Neoplasia, 2006. **8**(1): p. 69-78.

59. Rushmere, N.K., et al., *Analysis of the level of mRNA expression of the membrane regulators of complement, CD59, CD55 and CD46, in breast cancer.* International journal of cancer, 2003. **108**(6): p. 930-936.

60. Bjorge, L., et al., *Complement-regulatory proteins in ovarian malignancies.* International journal of cancer. Journal international du cancer, 1997. **70**(1): p. 14-25.

61. Kobayashi, D., et al., *Olfactomedin 4 promotes S-phase transition in proliferation of pancreatic cancer cells.* Cancer science, 2007. **98**(3): p. 334-40.

62. Kim, H., et al., *The retinoic acid synthesis gene ALDH1a2 is a candidate tumor suppressor in prostate cancer.* Cancer research, 2005. **65**(18): p. 8118-24.

63. Glavinas, H., et al., *The role of ABC transporters in drug resistance, metabolism and toxicity.* Current drug delivery, 2004. **1**(1): p. 27-42.

64. Zhou, L., et al., *Angiotensin-converting enzyme 2 acts as a potential molecular target for pancreatic cancer therapy.* Cancer letters, 2011. **307**(1): p. 18-25.

65. Ishizuka, J., C.M. Townsend, Jr., and J.C. Thompson, *Neurotensin regulates growth of human pancreatic cancer.* Annals of surgery, 1993. **217**(5): p. 439-45; discussion 446.

66. Ryder, N.M., et al., *G protein-coupled receptor signaling in human ductal pancreatic cancer cells: neurotensin responsiveness and mitogenic stimulation.* Journal of cellular physiology, 2001. **186**(1): p. 53-64.

67. Bernard-Pierrot, I., et al., *Characterization of the recurrent 8p11-12 amplicon identifies PPAPDC1B, a phosphatase protein, as a new therapeutic target in breast cancer.* Cancer research, 2008. **68**(17): p. 7165-75.

68. Kang, D., et al., *The histone methyltransferase Wolf-Hirschhorn syndrome candidate 1-like 1 (WHSC1L1) is involved in human carcinogenesis.* Genes, chromosomes & cancer, 2013. **52**(2): p. 126-39.

69. Erkan, M., et al., *The role of stroma in pancreatic cancer: diagnostic and therapeutic implications.* Nature reviews. Gastroenterology & hepatology, 2012. **9**(8): p. 454-67.

70. Erkan, M., et al., *Organ-, inflammation- and cancer specific transcriptional fingerprints of pancreatic and hepatic stellate cells.* Molecular cancer, 2010. **9**: p. 88.

71. Harada, T., et al., *Genome-wide analysis of pancreatic cancer using microarray-based techniques.* Pancreatology : official journal of the International Association of Pancreatology, 2009. **9**(1-2): p. 13-24.

72. Van den Broeck, A., et al., *Molecular markers associated with outcome and metastasis in human pancreatic cancer.* Journal of experimental & clinical cancer research : CR, 2012. **31**: p. 68.

73. Grutzmann, R., et al., *Gene expression profiles of microdissected pancreatic ductal adenocarcinoma.* Virchows Archiv : an international journal of pathology, 2003. **443**(4): p. 508-17.

74. Ivanov, S., et al., *Expression of hypoxia-inducible cell-surface transmembrane carbonic anhydrases in human cancer.* The American journal of pathology, 2001. **158**(3): p. 905-19.

75. Mochizuki, S. and Y. Okada, *ADAMs in cancer cell proliferation and progression.* Cancer science, 2007. **98**(5): p. 621-8.

76. Rocks, N., et al., *Emerging roles of ADAM and ADAMTS metalloproteinases in cancer.* Biochimie, 2008. **90**(2): p. 369-79.

77. Yuan, G., et al., *Oncogenic function of DACT1 in colon cancer through the regulation of beta-catenin.* PloS one, 2012. **7**(3): p. e34004.

78. Lehmann, B.D., et al., *Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies.* The Journal of clinical investigation, 2011. **121**(7): p. 2750-67.

79. Martoglio, A.M., et al., *Changes in tumorigenesis- and angiogenesis-related gene transcript abundance profiles in ovarian cancer detected by tailored high density cDNA arrays.* Molecular medicine, 2000. **6**(9): p. 750-65.

80. Wiesmann, F., et al., *Frequent loss of endothelin-3 (EDN3) expression due to epigenetic inactivation in human breast cancer.* Breast cancer research : BCR, 2009. **11**(3): p. R34.

81. Wang, Y., et al., *Identifying novel prostate cancer associated pathways based on integrative microarray data analysis.* Computational biology and chemistry, 2011. **35**(3): p. 151-8.

82. Muller, M.R. and A. Rao, *NFAT, immunity and cancer: a transcription factor comes of age.* Nature reviews. Immunology, 2010. **10**(9): p. 645-56.

83. Arrick, B.A. and C.F. Nathan, *Glutathione metabolism as a determinant of therapeutic efficacy: a review.* Cancer research, 1984. **44**(10): p. 4224-32.

84. Bolos, V., J. Grego-Bessa, and J.L. de la Pompa, *Notch signaling in development and cancer.* Endocrine reviews, 2007. **28**(3): p. 339-63.

85. Kastan, M.B. and J. Bartek, *Cell-cycle checkpoints and cancer.* Nature, 2004. **432**(7015): p. 316-23.

86. Yilmaz, M. and G. Christofori, *EMT, the cytoskeleton, and cancer cell invasion.* Cancer metastasis reviews, 2009. **28**(1-2): p. 15-33.

87. Hirohashi, S. and Y. Kanai, *Cell adhesion system and human cancer morphogenesis.* Cancer science, 2003. **94**(7): p. 575-81.

88. Tahara, H. and M.T. Lotze, *Antitumor effects of interleukin-12 (IL-12): applications for the immunotherapy and gene therapy of cancer.* Gene therapy, 1995. **2**(2): p. 96-106.

89. Colombo, M.P. and G. Trinchieri, *Interleukin-12 in anti-tumor immunity and immunotherapy.* Cytokine & growth factor reviews, 2002. **13**(2): p. 155-68.

90. Reya, T. and H. Clevers, *Wnt signalling in stem cells and cancer.* Nature, 2005. **434**(7035): p. 843-50.

91. Kirikoshi, H., H. Sekihara, and M. Katoh, *Up-regulation of Frizzled-7 (FZD7) in human gastric cancer.* International journal of oncology, 2001. **19**(1): p. 111-5.

92. Ueno, K., et al., *Frizzled-7 as a potential therapeutic target in colorectal cancer.* Neoplasia, 2008. **10**(7): p. 697-705.

93. Clarke, G.M., et al., *Basic statistical analysis in genetic case-control studies.* Nature protocols, 2011. **6**(2): p. 121-33.

94. Williams, S.M., et al., *Problems with genome-wide association studies.* Science, 2007. **316**(5833): p. 1840-2.

95. Andrianifahanana, M., et al., *Mucin (MUC) gene expression in human pancreatic adenocarcinoma and chronic pancreatitis: a potential role of MUC4 as a tumor marker of diagnostic significance.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2001. **7**(12): p. 4033-40.

96. Antonarakis, E.S., et al., *Phase I rapid dose-escalation study of AGS-1C4D4, a human anti-PSCA (prostate stem cell antigen) monoclonal antibody, in patients with castration-resistant prostate cancer: a PCCTC trial.* Cancer chemotherapy and pharmacology, 2012. **69**(3): p. 763-71.

97. Wolpin BM, O.R.E., Ko Y, *Global, multicenter, open-label, randomized phase II trial comparing gemcitabine (G) with G plus AGS-1C4D4 (A) in patients (pts) with metastatic pancreatic cancer (mPC).* J Clin Oncol., 2011. **29 (Suppl)**(abstract 4031).

98. Lu, X., et al., *ADAM proteins - therapeutic potential in cancer.* Current cancer drug targets, 2008. **8**(8): p. 720-32.

CHAPTER 2

MOLECULAR PROFILING PREDICTS THE EXISTENCE OF TWO FUNCTIONALLY DISTINCT CLASSES OF OVARIAN CANCER STROMA

2.1 Introduction

The epithelial cells of the ovary interact with the cells of the surrounding microenvironment in order to regulate tissue homeostasis. Morphologically, the normal ovarian epithelial cells form a flat-to-cuboidal monolayer supported by a basement membrane. Cells located below this basement membrane are composed of various cell types collectively referred to as stromal cells. The most common types of stromal cells are fibroblasts, pericytes, endothelial cells, and various immune and inflammatory cells. Stromal and epithelial cells communicate through the secretion and binding of growth factors and other signaling molecules that promote reciprocal cellular responses appropriate for coordinated cell functions, for example, those required for the replication of ovarian surface epithelial cells following ovulation [1-3].

During cancer progression, genetic and epigenetic alterations lead to changes in the morphology and behavior of both epithelial and stromal cells by disrupting the tissue architecture and by interfering with signaling mechanisms. For example, signaling changes in a wide variety of developing cancer cells have been shown to result in the disruption of tissue homeostasis by inducing extracellular matrix (ECM) turnover, basement membrane disassociation, and increased stromal cell proliferation [1, 4].

Despite the well-documented role of stromal cell signaling in cancer progression, relatively few studies have been focused specifically on epithelial ovarian cancer-stromal interactions (EOC-SI). Previously reported studies on EOC-SI have focused on particular stromal components [5, 6], on specific cell lines [7] and/or employed in-house fabricated

microarrays of limited scope [8]. We report here the results of a study of EOC-SI using high-throughput gene expression (microarray) analysis of normal ovarian surface epithelial cells and cells captured from normal stroma, cancer epithelia, and cancer stroma using laser capture microdissection (LCM). Our results reveal the existence of two categories of ovarian cancer stroma. Analysis of ligand-receptor patterns of gene expression indicates that one of these classes of cancer stroma may be more permissive and one more resistant to associated cancer cell growth.

2.2 Materials and Methods

2.2.1 Tissue collection

Tissues were collected at Northside Hospital (Atlanta, GA) under appropriate Institutional Review Board protocols. Following resection, the tumor tissues were placed in cryotubes and immediately (<1 minute) frozen in liquid nitrogen. Samples were transported on dry ice to Georgia Institute of Technology (Atlanta, GA), and stored at -80 ºC. All tissues were examined and diagnoses made by a pathologist. The histopathology for each sample is listed in Table 2.1.

Table 2.1. Patient samples used in this study

Patient ID	Age at time of surgery	Tissue for Microarray	Histopathology	Stage	Grade
460	65	OSE	WNL(Within Normal Limits)	N/A	N/A
552	41	OSE	WNL	N/A	N/A
563	66	OSE	WNL	N/A	N/A
567	78	OSE	WNL	N/A	N/A
434	41	OSE/NS	WNL	N/A	N/A
437	54	OSE/NS	WNL	N/A	N/A
440	50	OSE/NS	WNL	N/A	N/A
448	63	OSE/NS	WNL	N/A	N/A

"Table continued"

452	51	OSE/NS	WNL	N/A	N/A
463	48	OSE/NS	WNL	N/A	N/A
470	44	OSE/NS	WNL	N/A	N/A
475	63	OSE/NS	WNL	N/A	N/A
317	59	Cepi	serous adenocarcinoma	Ic	3
489	48	Cepi	serous adenocarcinoma	IV	3
528	66	Cepi	serous adenocarcinoma	IIIc	3
537	64	Cepi	serous adenocarcinoma	IIIa	2
542	61	Cepi	serous adenocarcinoma	IV	3
551	59	Cepi	serous adenocarcinoma	IIIc/IV	3
588	71	Cepi	serous adenocarcinoma	IIIc	2
606	54	Cepi	serous adenocarcinoma	IIIa	3
617	64	Cepi	serous adenocarcinoma	IIIc	2
620	62	Cepi	serous adenocarcinoma	III/IV	3
651	46	Cepi	serous adenocarcinoma	IIIb/IIIc	3
183	66	Cepi/CS	serous adenocarcinoma	III	2
369	52	Cepi/CS	serous adenocarcinoma	IIIc	2
229	58	Cepi/CS	serous adenocarcinoma	IIIc	3
242	63	Cepi/CS	serous adenocarcinoma	IIIb	3
336	63	Cepi/CS	serous adenocarcinoma	Ic	3
367	56	Cepi/CS	serous adenocarcinoma	II	3
413	49	Cepi/CS	serous adenocarcinoma	IIb	3

For each of the cancer tissue samples, 7μm frozen sections were cut from samples embedded in cryomatrix (Shandon), and attached to uncharged microscope slides. Immediately following dehydration and staining (HistoGene, LCM Frozen Section Staining Kit, Arcturus), slides were processed in an Autopix (Arcturus) instrument for

laser capture microdissection (LCM) of cancer epithelial cells (Cepi), cancer stroma (CS), and normal stroma (NS) using CapSure Macro-LCM Caps (Arcturus).

Approximately 30,000 cells were collected from each of the samples. Normal ovarian surface epithelial (OSE) cells were also collected from normal ovaries at the time of surgery by light brushing using a Cytobrush Plus (Medscand), immediately stabilized in RNAlater (Ambion), and subsequently stored at -20 ºC. Microscopic examination of all collected cells was carried out to confirm the integrity and purity of the samples.

2.2.2 RNA extraction and amplification

PicoPure RNA Isolation Kit (Arcturus) protocols were followed for RNA extraction from the LCM cells on the Macro-LCM caps in 25µL of extraction buffer. Normal OSE cells were pelleted from RNAlater, RNA was isolated with Trizol (Invitrogen), and purified with the PicoPure RNA Isolation Kit. RNA quality was verified for all samples on the Bioanalyzer RNA Pico Chip (Agilent Technologies).

Total RNA from the above extractions was processed using the RiboAmp HS kit (Arcturus) in conjunction with the IVT Labeling Kit from Affymetrix, to produce an amplified, biotin-labeled mRNA suitable for hybridizing to GeneChip Human Genome U133 Plus 2.0 Arrays (Affymetrix) following manufacturer's recommendations.

2.2.3 Microarray data analysis

We generated 45 individual gene expression profiles from 12 OSE brushings and 18 Cepi, 8 NS and 7 CS patient samples isolated by laser capture microdissection (LCM). Affymetrix .CEL files were processed using the Affymetrix Expression Console (EC) Software Version 1.1 with the default MAS5.0 probeset normalization algorithm. The expression values from the 12 OSE, 18 Cepi, 8 NS and 7 CS samples were $\log_2$ transformed and then averaged for each probeset across each sample type. The microarray data were deposited in the Gene Expression Omnibus (GSE38666).

Probesets (genes) with nearly constant expression values (log_2 normalized) across samples (SD < 1) were excluded from further consideration. Of the 54,675 probesets on the U133 Plus 2.0 chip, 42,698 were thus retained. A four-way ANOVA was subsequently employed to identify genes significantly differentially expressed ($p \leq 0.001$) across the four sample groups (OSE, Cepi, NS and CS). These 6,654 genes were employed in the initial clustering analysis.

A subsequent comparison among the CS samples (CS_1 & CS_2) alone was performed using a similar approach. Specifically, genes with nearly identical expression values (SD<1) across CS_1 & CS_2 were discarded and the remaining 38,972 genes were subjected to an unpaired t-test to identify those genes that were significantly differentially expressed between the CS_1 & CS_2 sub-groups ($p \leq 0.001$, 88 genes).

All heat maps were generated using the UPGMA (unweighted average) clustering method and the Euclidean distance similarity measure.

2.2.4 Ligand-receptor compatibility analysis

For the ligand-receptor compatibility analysis, probesets associated with no or marginal expression across all 45 samples were discarded resulting in 5,865 differentially expressed genes. The presence or absence of the expression in samples was determined using the Affymetrix default MAS 5.0 decision algorithm. The MAS 5.0 algorithm uses the Tukey's biweight estimator to provide a robust mean signal value and the Wilcoxon's rank test to calculate the significance of the signal or p-value and detection call (present, marginal or absent) for each probeset. The p-values upon which the presence-absence calls for each ligand and receptor is based are presented in the appropriate tables (below).

2.3 Results

2.3.1 Hierarchical clustering establishes two distinct classes of stroma among the ovarian cancer patient samples

Forty-five gene expression profiles were generated from 12 OSE brushings and 18 Cepi, 8 NS and 7 CS patient samples isolated by laser capture microdissection (LCM). The relevant histopathologies of these 45 samples are listed in Table 2.1. Expression analysis yielded 6,654 differentially expressed probesets among the four sample types (ANOVA, $p \leq 0.001$). Hierarchical clustering of these expression data resulted in clear separations between the OSE, Cepi, NS and CS samples (Figure 2.1). Interestingly, the CS samples sub-divided into two distinct groups. One (CS_1) was more closely associated with the NS samples and the other (CS_2) was more closely associated with the Cepi samples.

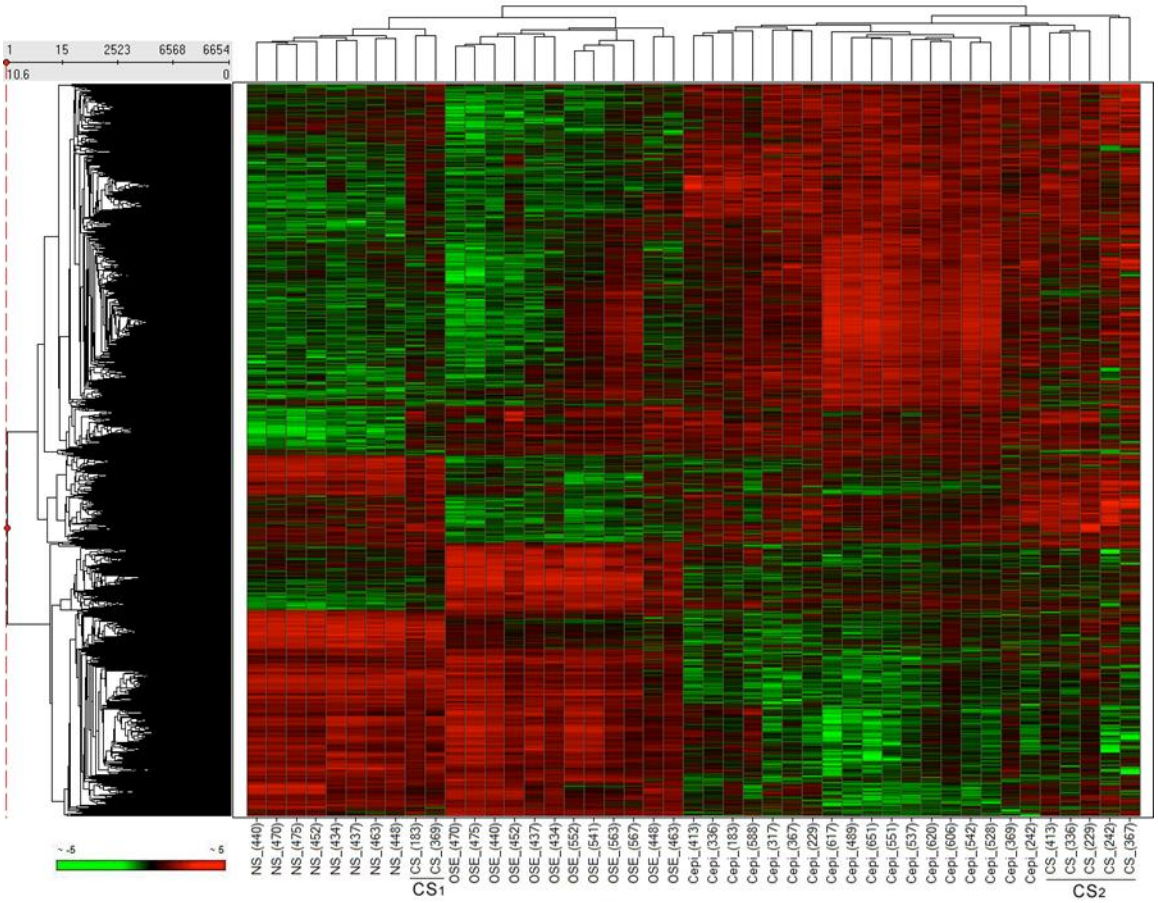


Figure 2.1. Hierarchical clustering of OSE, Cepi, NS and CS expression profiles. The heat map was generated by Z-score normalization of $\log_2$ expression values from Affymetrix HG U133 Plus 2.0. The results show that the OSE, Cepi, NS, and CS samples cluster into separate groups. The CS samples clustered into two distinct sub-groups (CS_1 and CS_2).

One possibility is that the two sub-classes of CS are simply a reflection of differential responses of stroma to molecular differences in the adjacent Cepi. If this were the case, we would expect to see a correlated sub-structure among the molecular profiles of the Cepi samples associated with the CS_1 and CS_2 sub-groups. As shown in Figure 2.1, no such coordinated sub-structure pattern exists among the Cepi samples indicating that the two sub-classes of CS are not merely a reflection of differential responses of the stroma to different Cepi sub-types.

As stated above, microscopic examination of LCM collected cells was carried out to validate the integrity of our samples. As a further confirmation, we conducted an additional computational analysis. In this analysis, probesets associated with no or marginal expression across all 45 samples were discarded resulting in 5,865 differentially expressed genes. If the reason for the presence of two classes of CS samples is that the CS_2 class was a mixture of stroma and invasive Cepi cells, the gene expression levels in the putative mixed cancer stromal class (CS_2) would be expected to lie within the range of the maximum and minimum expression levels of the NS and Cepi groups (*i.e.*, $avg(CS_2) < Min$ and $avg(CS_2) > Max$). Inconsistent with this prediction, we found that 2,342 or 40% (2,342/5,865) of the differentially expressed genes making up the CS_2 class displayed values outside the predicted range of the mixed cell types. The fact that 60% of the expression values lie within the predicted range is not indicative of contamination but rather of the fact that not all genes are significantly over-expressed in the stroma relative to cancer samples. Collectively our microarray results are consistent with the microscopic examination in demonstrating the absence of infiltrating Cepi cells in the cancer stroma samples.

2.3.2 Gene expression patterns are consistent with the existence of ligand-receptor interactions between Cepi and CS

The significance of the presence of two distinct classes of ovarian cancer stroma may involve differential interaction between these stroma and the adjacent cancer cells. To explore this possibility, we first examined the expression levels of genes encoding signaling ligands and compatible receptors in the CS and Cepi datasets.

Two lists were established from the 5,865 differentially expressed probesets across the OSE, Cepi, NS and CS samples. One list is comprised of all differentially expressed gene probes (note that each gene may be represented by multiple, non-overlapping probes) encoding secreted ligands (ligand-list), and the other all expressed gene probes encoding surface receptors (receptor-list) with documented binding affinity to the differentially expressed ligands (compatible ligands and receptors). The ligand-list consists of 34 CS and 36 Cepi ligands while the receptor-list is comprised of 20 Cepi and 21 CS receptors (Tables 2.2a and 2.2b).

Table 2.2. The 34 CS expressed ligands with the 20 expressed Cepi receptors (2a) and the 36 Cepi expressed ligands with the 21 expressed CS receptors (2b) (Significance of detection calls: *p ≤ 0.05, **p ≤ 0.005, *** p ≤ 0.0005).

(a)

CS Ligands		Cepi Receptors	
Gene Symbol	**Probeset ID**	**Gene Symbol**	**Probeset ID**
***CXCL1	204470_at	****CXCR4	217028_at
*CXCL3	207850_at	**FGFR2	208228_s_at
***CXCL9	203915_at	***FGFR3	204379_s_at
***CXCL10	204533_at	**MET	203510_at
***CXCL11	210163_at	*TGFBR2	207334_s_at
***CXCL12	209687_at	***TGFBR2	208944_at
***CXCL12	203666_at	**TGFBR3	204731_at
***CXCL13	205242_at	***TGFBR3	226625_at
**CXCL16	223454_at	***PDGFRA	203131_at
CXCL17	226960_at	*PDGFRA	1554828_at
*FGF1	205117_at	***IL1R1	202948_at
*FGF2	204422_s_at	*IL1R1	215561_s_at
***FGF7	1554741_s_at	*IL1R2	205403_at
*FGF9	239178_at	**IL7R	226218_at

(b)

Cepi Ligands		CS Receptors	
Gene Symbol	**Probeset ID**	**Gene Symbol**	**Probeset ID**
***CXCL1	204470_at	****CXCR4	217028_at
**CXCL3	207850_at	**FGFR2	208228_s_at
***CXCL9	203915_at	*FGFR3	204379_s_at
***CXCL10	204533_at	IL12RB1	1552584_at
***CXCL11	210163_at	***IL1R1	202948_at
*CXCL12	203666_at	***TGFBR2	208944_at
*CXCL12	209687_at	***TGFBR3	204731_at
**CXCL13	205242_at	***TGFBR3	226625_at
**CXCL16	223454_at	***PDGFRA	203131_at
CXCL17	226960_at	**PDGFRA	215305_at
*FGF1	205117_at	***MET	203510_at
***FGF9	206404_at	**IL1R1	215561_s_at
***FGF9	239178_at	*IL1R2	205403_at
*FGF11	227271_at	**IL7R	226218_at

"Table continued"

**FGF9	206404_at	**IL10RA	204912_at	***FGF18	231382_at	**IL10RA	204912_at
***FGF13	205110_s_at	**FZD1	204451_at	*FGF18	211029_x_at	*IL21R	221658_s_at
*HGF	210997_at	**FZD2	210220_at	*FGF18	206987_x_at	**FZD1	204451_at
***IGF1	209540_at	**FZD7	203705_s_at	**FGF18	214284_s_at	**FZD2	210220_at
*IGF2	202409_at	***FZD7	203706_s_at	**TGFA	205016_at	***FZD7	203705_s_at
*TGFA	205016_at	**FZD10	219764_at	**TGFB2	209909_s_at	**FZD7	203706_s_at
***TGFB2	209909_s_at			**PDGFA	205463_s_at	***FZD10	219764_at
*PDGFA	205463_s_at			***PDGFD	219304_s_at		
***PDGFD	219304_s_at			***IGF1	209540_at		
*IL7	206693_at			*IL7	206693_at		
***IL15	205992_s_at			*IL1B	39402_at		
**IL16	209828_s_at			***IL15	205992_s_at		
**IL17D	227401_at			**IL18	206295_at		
**IL18	206295_at			**WNT2	205648_at		
**WNT2B	206458_s_at			**WNT2B	206458_s_at		
*WNT7A	210248_at			***WNT5A	213425_at		
***WNT5A	213425_at			*WNT7A	210248_at		
***VEGFA	210512_s_at			**WNT11	206737_at		
*VEGFA	210513_s_at			***VEGFA	210512_s_at		
*VEGFA	211527_x_at			***VEGFA	210513_s_at		
				*VEGFA	211527_x_at		
				**VEGFA	212171_x_at		

We considered the expression of a ligand in CS (or Cepi) and its compatible receptor in Cepi (or CS) to be indicative of a potential CS-Cepi signaling interaction. Based on these criteria, we identified potential epithelial cancer-stroma signaling interactions (34 CS ligands and 20 Cepi receptors, see Table 2.2a, and 36 Cepi ligands and 21 CS receptors, see Table 2.2b). Of these, there were 17 compatible pairs for both the CS ligands-Cepi receptors and Cepi ligands-CS receptors interactions (Table 2.3).

Table 2.3. The expressed, compatible ligands and receptors as potential interactions between the Cepi and the CS samples from Tables 2.2a and 2.2b (significance of detection calls *p ≤ 0.05, **p ≤ 0.005, ***p ≤ 0.0005).

(a) **(b)**

CS Ligands	Probesets	Compatible Cepi Receptors	Probesets	Cepi Ligands	Probesets	Compatible CS Receptors	Probesets
***CXCL12	203666_at	***CXCR4	217028_at	*CXCL12	203666_at	***CXCR4	217028_at
***CXCL12	209687_at	***CXCR4	217028_at	*CXCL12	209687_at	***CXCR4	217028_at
*FGF1	205117_at	**FGFR2	208228_s_at	*FGF1	205117_at	**FGFR2	208228_s_at
*FGF1	205117_at	*FGFR3	204379_s_at	*FGF1	205117_at	*FGFR3	204379_s_at
*FGF2	204422_s_at	*FGFR3	204379_s_at	*FGF9	206404_at	*FGFR3	204379_s_at
**FGF9	206404_at	*FGFR3	204379_s_at	FGF9	239178_at	*FGFR3	204379_s_at
*FGF9	239178_at	*FGFR3	204379_s_at	*PDGFA	205463_s_at	***PDGFRA	203131_at
*HGF	210997_at	*MET	203510_at	*PDGFA	205463_s_at	*PDGFRA	215305_at
*PDGFA	205463_s_at	*PDGFRA	1554828_at	*TGFB2	209909_s_at	***TGFBR2	208944_at
*PDGFA	205463_s_at	***PDGFRA	203131_at	*WNT2	205648_at	*FZD2	210220_at
***TGFB2	209909_s_at	*TGFBR2	207334_s_at	*WNT2B	206458_s_at	***FZD10	219764_at
***TGFB2	209909_s_at	***TGFBR2	208944_at	*WNT7A	210248_at	***FZD7	203705_s_at
*WNT2	205648_at	*FZD2	210220_at	*WNT7A	210248_at	**FZD7	203706_s_at
**WNT2B	206458_s_at	*FZD10	219764_at	*IL1B	39402_at	***IL1R1	202948_at
*WNT7A	210248_at	*FZD7	203705_s_at	*IL1B	39402_at	*IL1R2	205403_at
*WNT7A	210248_at	**FZD7	203706_s_at	*IL1B	39402_at	*IL1R1	215561_s_at
*IL7	206693_at	**IL7R	226218_at	*IL7	206693_at	*IL7R	226218_at

Viewed from the perspective of individual genes (*i.e.*, combining multiple probes of the same genes) there were 12 unique CS ligand-Cepi receptor pairs, and 12 unique Cepi ligand-CS receptor pairs in our observed dataset (Table 2.4).

Table 2.4. The unique compatible ligands and receptors as potential interactions between the Cepi and the CS samples when multiple probes from Tables 2.3a and 2.3b are combined (significance of detection calls: *p ≤ 0.05, **p ≤ 0.005, ***p ≤ 0.0005).

CS Ligands	Compatible Cepi Receptors	Cepi Ligands	Compatible CS Receptors
***CXCL12	***CXCR4	*CXCL12	***CXCR4
*FGF1	**FGFR2	*FGF1	**FGFR2
*FGF1	*FGFR3	*FGF1	*FGFR3
*FGF2	*FGFR3	**FGF9	*FGFR3

"Table continued"

**FGF9	*FGFR3	**PDGFA	***PDGFRA
*HGF	*MET	**TGFB2	***TGFBR2
*PDGFA	***PDGFRA	**IL7	*IL7R
***TGFB2	***TGFBR2	**IL1B	*IL1R1
*IL7	**IL7R	**IL1B	*IL1R2
*WNT2	*FZD2	**WNT2	*FZD2
**WNT2B	*FZD10	**WNT2B	***FZD10
*WNT7A	**FZD7	**WNT7A	***FZD7

To determine if the observed co-expression of these 17 pairs of compatible ligands and receptors (probes) was greater than what would be expected by chance, we generated two lists. One list of observed data consisted of the expressed probes of the 17 CS ligands and 17 compatible Cepi receptors (Table 2.3a), and the other of the 17 expressed Cepi ligands and 17 CS receptors (Table 2.3b). A second list of random associations was generated using the same number of pairings as in the observed list (17 random pairs) and randomly selecting 17 pairs of ligands and receptors. One randomly selected CS (or Cepi) ligand from the pool of the 34 CS (or 36 Cepi) expressed ligands (Table 2.2a) was paired with one randomly selected Cepi (or CS) receptor from the pool of the 20 Cepi (or 21 CS) expressed receptors (Table 2.2b). These random associations were generated 100 times and each time the number of biologically compatible ligand-receptor pairs arising by chance was counted. The number of biologically compatible interactions in the observed data (17) was then compared to the number of compatible interactions scored from the randomized associations using Z-statistics. Two types of comparisons were performed, one for the pairs of CS ligands and Cepi receptors, and another for the pairs of Cepi ligands and CS receptors. For both comparisons, the observed number of biologically compatible ligand-receptor pairs was significantly greater than what is expected by chance (CS ligands-Cepi receptors z-score= -4.68, $p \leq 0.0002$; Cepi ligands-CS receptors z-score= -4.35, $p \leq 0.0002$). Thus, the observed co-expression of pairs of compatible ligands and receptors is biologically significant.

2.3.3 Specific ligand-receptor pairs between Cepi and CS show differential gene expression in the two CS classes

Of the 24 compatible pairs of ligand- and receptor-encoding genes listed in Table 2.4, most display similar expression patterns between CS_1 and CS_2. However, 6 of the ligand and receptor pairs display differential patterns of expression between the two groups of CS suggesting that CS_2 may be a more conducive microenvironment for tumor growth (Table 2.5). For example, the FGF2 ligand, a documented inhibitor of tumor growth [28], is expressed in NS and in CS_1 but not in CS_2. Since a compatible receptor of this inhibitor (FGFR3) is expressed in Cepi, CS_2 may be a more conducive microenvironment for tumor growth than CS_1. The Interleukin 7 (IL7) ligand has been previously implicated as an inducer of tumor growth in lymphoblastic leukemia [9], prostate cancer [10], breast cancer [11], and colorectal cancer [12]. IL7 is expressed in CS_2 but not in CS_1, again suggesting that CS_2 may be a more conducive microenvironment for tumor growth than CS_1.

Table 2.5. The unique compatible ligands and receptors from Table 2.4 showing the expression pattern in NS, CS_1, CS_2 and Cepi. The 6 highlighted signals had the same expression in NS and CS_1 but different expression between CS_1 and CS_2 despite the fact that their compatible signals in Cepi were always expressed. Expression is denoted with "+" (*i.e.*, there is at least one Affymetrix present call with detection p-value ≤ 0.05) and non-expression with "-" (*i.e.*, there are no Affymetrix present calls in the samples with detection p-value ≤ 0.05).

NS	LIGANDS	CS_1	CS_2	RECEPTORS	Cepi
+	CXCL12	+	+	CXCR4	+
-	FGF1	+	+	FGFR2	-
-	FGF1	+	+	FGFR3	+
+	FGF2	+	-	FGFR3	+
+	FGF9	+	+	FGFR3	+
-	HGF	+	+	MET	+
+	PDGFA	+	+	PDGFRA	+

"Table continued"

+	TGFB2	+	+	TGFBR2	+
-	IL7	-	+	IL7R	+
-	WNT2	+	+	FZD2	+
+	WNT2B	+	-	FZD2	+
-	WNT7A	-	+	FZD7	+
NS	RECEPTORS	CS_1	CS_2	LIGANDS	Cepi
+	CXCR4	+	+	CXCL12	+
+	FGFR2	+	+	FGF1	+
-	FGFR3	-	+	FGF1	+
-	FGFR3	-	+	FGF9	+
+	PDGFRA	+	+	PDGFA	+
+	TGBFR2	+	+	TGFB2	+
+	IL1R1	+	+	IL1B	+
-	IL1R2	+	-	IL1B	+
+	IL7R	+	+	IL7	+
+	FZD2	+	+	WNT2	+
+	FZD2	+	+	WNT2B	+
+	FZD7	+	+	WNT7A	+

The well-documented cancer inducing ligands FGF1 and FGF9 [13-15] are both highly expressed in Cepi. The fact that the compatible FGFR3 receptor is expressed in CS_2 but not in CS_1 again suggests that CS_2 is a more favorable microenvironment for ovarian cancer growth than CS_1

The *WNT* family of genes is involved in a variety of developmental processes and aberrant expression of various members of *WNT* genes have been implicated in cancer [16]. For example, WNT7A is a ligand present in the extracellular matrix that participates in the sexual development of the Mullerian ducts [17]. Recent *in vivo* mouse studies suggest that WNT7A is an inducer of ovarian cancer growth [18]. Consistent with this interpretation, WNT7A has recently been identified as a potential early stage biomarker of human ovarian cancer [19]. The fact that WNT7A is expressed in CS_2 but not in CS_1 is also consistent with the hypothesis that CS_2 may be a more conducive microenvironment for ovarian cancer growth than CS_1.

A second member of the WNT family, WNT2B, is expressed in CS1 but not CS2 suggesting, contrary to what is presented above, that CS_1 may be more permissive for cancer growth. However, the fact that WNT2B has been previously reported to be expressed in normal ovaries [20] coupled with our finding that it is also expressed in NS, makes interpreting the significance of the dichotomy in WNT2B expression between CS1 and CS2 ambiguous.

2.4 Discussion

Cancer progression is a dynamic process involving cellular adaptation and survival that is, in part, driven by signaling interactions between participating cells. Many signaling interactions have been documented to take place between cancer epithelial cells and the surrounding stroma [21]. Early in tumor development, cancer cells produce growth factors that are believed to modulate or "activate" the surrounding stroma in order to convert the stroma into a supportive microenvironment for cancer growth [2, 13]. For example, it has been shown that growth factors secreted by macrophages can contribute to cancer progression and metastasis [22]. Other inflammatory cells such as lymphocytes, neutrophils, mast cells, T-regulatory cells and platelets also have been shown to have the potential to support tumor progression by negatively regulating the anti-cancer host immune response [23-25]. Fibroblasts, the major component of the stroma, have been shown to be able to participate actively in the malignant progression of cancer by producing growth factors, various chemokines and extra cellular matrix components that facilitate the production of endothelial cells and pericytes conducive to tumor growth [13, 26].

The purpose of this study was to investigate the process of stroma activation within the context of ovarian cancer. Toward this end, we conducted RNA microarray profiling of 45 tissue samples using the Affymetrix (U133 Plus2) gene expression platform. Laser capture microdissection (LCM) was used to isolate cancer cells from the

tumors of 18 ovarian cancer patients (Cepi). For 7 of these patients, a matched set of surrounding cancer stroma (CS) was also collected. For controls, we isolated surface epithelial cells (OSE) from the normal (non-cancerous) ovaries of 12 individuals including matched sets of samples of OSE and normal stroma (NS) from 8 of these patients.

Unsupervised hierarchical clustering of the microarray data resulted in the expected separation between the OSE and Cepi samples. Consistent with models of stromal activation, we also observed significant separation between the NS and CS samples. Somewhat unexpected, however, was our finding that the CS samples clustered into two distinct sub-groups (CS_1 and CS_2).

Based on patterns of co-expression of ligand and receptor encoding genes, we determined that 6 biologically compatible pairs of ligands and receptors are differentially expressed between Cepi and the CS_1 and CS_2 cancer stroma. The patterns of differential expression between the compatible ligands and receptors are consistent with the hypothesis that CS_2 may be a more conducive microenvironment for tumor growth (Table 2.5). For example, the expression of tumor promoting ligands in Cepi is always matched with the expression of compatible receptors in CS_2 but not in CS_1.

The fact that certain tumor microenvironments are capable of inhibiting tumor growth and/or development is well established. For example, macrophages can act as anti-cancer agents within the context of the innate immune response [27]. Likewise, fibroblasts, in some cellular contexts, have been shown to revert tumor cells to a normal, non-cancerous phenotype [28, 29]. Normal ovarian stromal cells have been shown to significantly inhibit ovarian cancer cell growth when co-injected into nude mice [30].

The apparently innate anti-cancer properties of normal stroma are generally considered to be transient giving way to the “activation” of pro-cancer growth signals induced by cancer cells as the tumors progress [1]. However, since the majority of the patients associated with the CS_1 class of cancer stroma have, like the majority of the

cancer patients included in our study, already progressed to advanced staged disease (Table 2.1), it is unlikely that the CS_1 molecular profile represents a transient condition. Rather, our results suggest that variability exists among ovarian cancer patients with respect to the propensity of normal stroma to become activated.

2.5 Conclusions

An understanding of the potential clinical significance of the observed molecular dichotomy between ovarian cancer stroma is beyond the scope of this present study. However, it is relevant to note that all of the cancers associated with the putatively more permissive CS_2 cancer stroma were classified as grade 3 while those associated with the putatively more resistant CS_1 cancer stroma were classified as grade 2. The fact that no distinction was apparent between the molecular profiles of grade 2 and grade 3 Cepi samples (Figure 2.1), suggests that cancer grade may, at least in part, be determined by the relative permissiveness of the tumor microenvironment. Molecular profiling of larger numbers of matched sets of ovarian cancer and stroma samples will be required to adequately test this hypothesis. Nevertheless, the current results are consistent with the hypothesis that the microenvironment plays a significant role in ovarian cancer development and suggest that functionally significant variability may exist among ovarian cancer patients in the ability of the microenvironment to modulate cancer development.

2.6 References

1. Liotta, L.A. and E.C. Kohn, *The microenvironment of the tumour-host interface.* Nature, 2001. 411(6835): p. 375-9.

2. Bhowmick, N.A. and H.L. Moses, *Tumor-stroma interactions.* Curr Opin Genet Dev, 2005. 15(1): p. 97-101.

3. Tlsty, T.D. and L.M. Coussens, *Tumor stroma and regulation of cancer development.* Annu Rev Pathol, 2006. 1: p. 119-50.

4. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer.* Cell, 2000. 100(1): p. 57-70.

5. Kamat, A.A., et al., *The clinical relevance of stromal matrix metalloproteinase expression in ovarian cancer.* Clin Cancer Res, 2006. 12(6): p. 1707-14.

6. Zhang, L., et al., *Tumor-derived vascular endothelial growth factor up-regulates angiopoietin-2 in host endothelium and destabilizes host vasculature, supporting angiogenesis in ovarian cancer.* Cancer Res, 2003. 63(12): p. 3403-12.

7. Porcile, C., et al., *Stromal cell-derived factor-1alpha (SDF-1alpha/CXCL12) stimulates ovarian cancer cell growth through the EGF receptor transactivation.* Exp Cell Res, 2005. 308(2): p. 241-53.

8. Wang, E., et al., *Peritoneal and subperitoneal stroma may facilitate regional spread of ovarian cancer.* Clin Cancer Res, 2005. 11(1): p. 113-22.

9. Silva, A., et al., *Intracellular reactive oxygen species are essential for PI3K/Akt/mTOR-dependent IL-7-mediated viability of T-cell acute lymphoblastic leukemia cells.* Leukemia : official journal of the Leukemia Society of America, Leukemia Research Fund, U.K, 2011. 25(6): p. 960-7.

10. Schroten, C., et al., *The additional value of TGFbeta1 and IL-7 to predict the course of prostate cancer progression.* Cancer immunology, immunotherapy : CII, 2011.

11. Al-Rawi, M.A., R.E. Mansel, and W.G. Jiang, *Interleukin-7 (IL-7) and IL-7 receptor (IL-7R) signalling complex in human solid tumours.* Histology and histopathology, 2003. 18(3): p. 911-23.

12. Maeurer, M.J., et al., *Interleukin-7 (IL-7) in colorectal cancer: IL-7 is produced by tissues from colorectal cancer and promotes preferential expansion of tumour infiltrating lymphocytes.* Scandinavian journal of immunology, 1997. 45(2): p. 182-92.

13. Bhowmick, N.A., E.G. Neilson, and H.L. Moses, *Stromal fibroblasts in cancer initiation and progression.* Nature, 2004. 432(7015): p. 332-7.

14. Jin, C., et al., *Directionally specific paracrine communication mediated by epithelial FGF9 to stromal FGFR3 in two-compartment premalignant prostate tumors.* Cancer research, 2004. 64(13): p. 4555-62.

15. Turner, N. and R. Grose, *Fibroblast growth factor signalling: from development to cancer.* Nature reviews. Cancer, 2010. 10(2): p. 116-29.

16. Reya, T. and H. Clevers, *Wnt signalling in stem cells and cancer.* Nature, 2005. 434(7035): p. 843-50.

17. Bui, T.D., et al., *Isolation of a full-length human WNT7A gene implicated in limb development and cell transformation, and mapping to chromosome 3p25.* Gene, 1997. 189(1): p. 25-9.

18. Yoshioka, S., et al., *WNT7A Regulates Tumor Growth and Progression in Ovarian Cancer through the WNT/beta-Catenin Pathway.* Molecular cancer research : MCR, 2012. 10(3): p. 469-82.

19. Tchagang, A.B., et al., *Early detection of ovarian cancer using group biomarkers.* Molecular cancer therapeutics, 2008. 7(1): p. 27-37.

20. Ricken, A., et al., *Wnt signaling in the ovary: identification and compartmentalized expression of wnt-2, wnt-2b, and frizzled-4 mRNAs.* Endocrinology, 2002. 143(7): p. 2741-9.

21. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. 144(5): p. 646-74.

22. Nowicki, A., et al., *Impaired tumor growth in colony-stimulating factor 1 (CSF-1)-deficient, macrophage-deficient op/op mouse: evidence for a role of CSF-1-dependent macrophages in formation of tumor stroma.* International journal of cancer. Journal international du cancer, 1996. 65(1): p. 112-9.

23. Pawelec, G., *Tumour escape: antitumour effectors too much of a good thing?* Cancer immunology, immunotherapy : CII, 2004. 53(3): p. 262-74.

24. Yang, L. and D.P. Carbone, *Tumor-host immune interactions and dendritic cell dysfunction.* Advances in cancer research, 2004. 92: p. 13-27.

25. Coussens, L.M. and Z. Werb, *Inflammation and cancer.* Nature, 2002. 420(6917): p. 860-7.

26. Kalluri, R. and M. Zeisberg, *Fibroblasts in cancer.* Nat Rev Cancer, 2006. 6(5): p. 392-401.

27. Gillessen, S., et al., *CD1d-restricted T cells regulate dendritic cell function and antitumor immunity in a granulocyte-macrophage colony-stimulating factor-dependent fashion.* Proceedings of the National Academy of Sciences of the United States of America, 2003. 100(15): p. 8874-9.

28. Micke, P. and A. Ostman, *Tumour-stroma interaction: cancer-associated fibroblasts as novel targets in anti-cancer therapy?* Lung cancer, 2004. 45 Suppl 2: p. S163-75.

29. Gonda, T.A., et al., *Molecular biology of cancer-associated fibroblasts: can these cells be targeted in anti-cancer therapy?* Seminars in cell & developmental biology, 2010. 21(1): p. 2-10.

30. Parrott, J.A., et al., *Stromal-epithelial interactions in the progression of ovarian cancer: influence and source of tumor stromal cells.* Molecular and cellular endocrinology, 2001. 175(1-2): p. 29-39.

CHAPTER 3

MOLECULAR PROFILING SUPPORTS THE ROLE OF EPITHELIAL-TO-MESENCHYMAL TRANSITION (EMT) IN OVARIAN CANCER METASTASIS

3.1 Introduction

While metastasis ranks among the most lethal of all cancer-associated processes, on the molecular level, it remains one of the least well understood [1]. One model that has gained credibility in recent years is that metastasizing cells at least partially recapitulate the developmental process of epithelial-to-mesenchymal transition (EMT) in their transit from primary to metastatic sites [2, 3]. While experimentally supported by cell culture and animal model studies (e.g., [4-7]), the lack of unambiguous confirmatory evidence in cancer patients has led to persistent challenges to the model's relevance in humans [8, 9]. We report here the results of gene expression profiling of 14 matched sets of primary and metastatic ovarian cancer (serous adenocarcinoma) patient samples.

While histological examination reveled no morphological distinction between matched sets of primary and metastatic samples, gene expression profiling clearly distinguished two classes of metastatic samples. One class displayed expression patterns statistically indistinguishable from primary samples isolated from the same patients while a second class displayed expression patterns significantly different from primary samples. Further analyses focusing on genes previously associated with EMT clearly distinguished the primary from metastatic samples in all but one patient. Our results are consistent with a role of EMT in most if not all ovarian cancer metastases and demonstrate that identical

morphologies between primary and metastatic cancer samples is insufficient evidence to negate a role of EMT in the metastatic process.

3.2 Materials and Methods

3.2.1 Tissue collection

Tissues were collected at Northside Hospital (Atlanta, GA) under appropriate Institutional Review Board protocols. Following resection, the tumor tissues were placed in cryotubes and immediately (<1 minute) frozen in liquid nitrogen. Samples were transported on dry ice to Georgia Institute of Technology (Atlanta, GA), and stored at -80 ºC. After examination and verification by a pathologist, tissues were embedded in cryomatrix (Shandon). The clinical information of the primary and metastatic cancer tissues from the seven patients is found in Figure 3.1).

Patient ID	Age	Histopathology	Stage	Grade	Morphological Comparison Primary vs Metastasis
489	48	Serous Adenocarcinoma	IV	3	Right ovary vs. Omentum: Similar morphology
528	66	Serous Adenocarcinoma	IIIc	3	Right ovary vs. Omentum: Similar morphology
542	61	Serous Adenocarcinoma	IV	3	Left ovary vs Omentum: Similar morphology
551	59	Serous Adenocarcinoma	IIIc/IV	3	Right ovary vs. omentum: Similar morphology
588	71	Serous Adenocarcinoma	IIIc	2/3	Right ovary vs. Omentum: Similar morphology
617	64	Serous Adenocarcinoma	IIIc	2/3	Left ovary vs. omentum: Similar morphology
620	62	Serous Adenocarcinoma	III/IV	3	Left ovary vs. omentum: Similar morphology

Figure 3.1 (a) Tissues from the primary and metastatic samples of the patients in the study display indistinguishable morphologies. **(b)** Clinical information of the patients in the study.

For each of the primary and metastatic (omental) tissue samples, 7μm frozen sections were cut and attached to uncharged microscope slides. Immediately following dehydration and staining (HistoGene, LCM Frozen Section Staining Kit, Arcturus. Life Technologies), slides were processed in an Autopix (Arcturus) instrument for laser capture microdissection (LCM). CapSure Macro-LCM Caps (Arcturus, Life

Technologies) were used to ensure purity of all primary cancer epithelial cells and omental metastatic cells. Approximately 30,000 cells were collected for each of the 14 tissue samples (seven primary and seven matched metastatic ovarian cancer samples).

3.2.2 RNA extraction and amplification

PicoPure RNA Isolation Kit (Life Technologies) protocols were followed for RNA extraction from the cells on the Macro-LCM caps in 30µL of extraction buffer. RNA quality was verified for all samples on the Bioanalyzer RNA Pico Chip (Agilent Technologies). Total RNA from the above extractions was processed using the Ovation® Pico WTA System (NuGEN) in conjunction with the Encore™ BiotinIL Module (NuGEN), to produce an amplified, biotin-labeled cDNA suitable for hybridizing to GeneChip Human Genome U133 Plus 2.0 Arrays (Affymetrix) following manufacturer's recommendations.

3.2.3 Microarray data analysis

Gene analysis: Fourteen individual gene expression profiles were generated from the primary and matched metastatic samples of each of the seven patients used in this study. The 14 Affymetrix .CEL files were processed using the Affymetrix Expression Console (EC) Software Version 1.1 using the Robust Multi-Array Average (RMA) normalization method. The normalized expression values from all 14 samples were $\log_2$ transformed. From the initial $\log_2$ transformed 54,675 probe sets (21,049 probe sets transformed, unique, annotated genes), 50,286 that displayed marginal differences in expression across all patient samples (standard deviation ≤ 0.8 from the mean of all 14 samples) were filtered out. The remaining 4,389 probe sets (3,365 genes) were employed in the unsupervised clustering analysis (Figure 3.2).

For the identification of the epithelial to mesenchymal transition (EMT) related genes, a list of 84 genes, previously implicated in the process of EMT, was employed

(Table A.3.2 (http://www.sabiosciences.com/rt_pcr_product/HTML/PAHS-090Z.html). Of these genes, 39 (61 probe sets) were identified among the 3,365 genes (4,389 probe sets) differentially expressed across samples (Figure 3.2, Table A.3.1).

Individual clustering analyses were carried out for matched sets of primary and metastatic samples from the Group 1 (Figure 3.2b) and Group 2 (Figure 3.2c) patients.

Gene clustering metrics: All hierarchical clustering (Figure 3.2, Figure 3.3) was performed using normalized Z-scores of the $\log_2$ transformed expression values.

Functional pathway enrichment: Biological interpretations of the differential gene expression data were performed by pathway enrichment analysis using MetaCore 5.2 (GeneGO, St Joseph, MI, http://thomsonreuters.com/products_services/science/systems-biology/).

3.3 Results/Discussion

EMT is a process by which epithelial cells acquire mesenchymal cell characteristics including reduced cell adhesiveness and increased cell motility [10]. The process is an essential component of embryonic development and is known to be both transient and reversible (mesenchymal-to-epithelial transition or MET). Initial observations that many characteristics of cancer metastasis appear to be recapitulations of key features of EMT/MET, led to the hypothesis that similar, if not identical, molecular mechanisms may be involved [11]. While numerous subsequent studies conducted in cell lines and animal models have supported this hypothesis (e.g., [4-7]), the lack of unambiguous evidence from studies involving human tumor samples has resulted in skepticism [8, 9]. One persistent objection to the model is the fact that careful morphological examinations of human metastases have never uncovered the existence of cancer mesenchymal-like cells [8]. Indeed, while it is clear that cancer metastasis must involve detachment of cancer cells from the primary tumor, such a phenomenon could well be attributed to mutation(s) and/or aberrant gene expression patterns in one or a few

genes and not necessarily reflect involvement of a more coordinated process such as EMT [8]. Although these alternative possibilities are not mutually exclusive, a clarification of the molecular mechanism(s) underlying metastasis is critical because it could have significant ramifications on future directions in the development of diagnostic tests and potential anti-metastatic drugs [12].

Ovarian cancer is the most malignant of all gynecological cancers and is responsible for over 14,000 deaths per year in the United States alone [13]. Because ovarian cancer is often asymptomatic early in its progression (Stages I/II), the disease is not typically diagnosed until later stages (Stages III/IV) when the cancer has metastasized and prognosis is poor (5 year survival < 30%) [13]. Ovarian cancer metastasis is primarily due to the exfoliation of malignant cells or cell aggregates from the primary tumor into the abdominal cavity and their subsequent spread and attachment to visceral and parietal peritoneal surfaces of abdominal organs such as the omentum. This mechanism of intra-abdominal metastatic spread allows for the capture and molecular comparison of primary and metastatic cancer cells isolated from the same patient, providing a favorable opportunity to evaluate the potential role of epithelial-to-mesenchymal transition (EMT) in the metastatic process.

Fourteen matched sets of primary and metastatic (omental) samples were collected from seven advanced staged (III/IV) ovarian cancer (serous adenocarcinoma) patients. Pathological examination classified all of the cancer samples as highly un-differentiated with no significant difference in morphology between any of the primary and metastatic sets (Figure 3.2). Tissue samples for molecular analysis were snap frozen in liquid nitrogen within one minute of surgical removal and subsequently embedded in cryomatrix (Shandon) for frozen sectioning and laser capture microdissection (LCM). Approximately 30,000 cancer epithelial cells were isolated from each tissue sample and RNA was extracted and processed for gene expression analysis (Affymetrix, U133Plus 2.0 arrays) according to previously described methods [14].

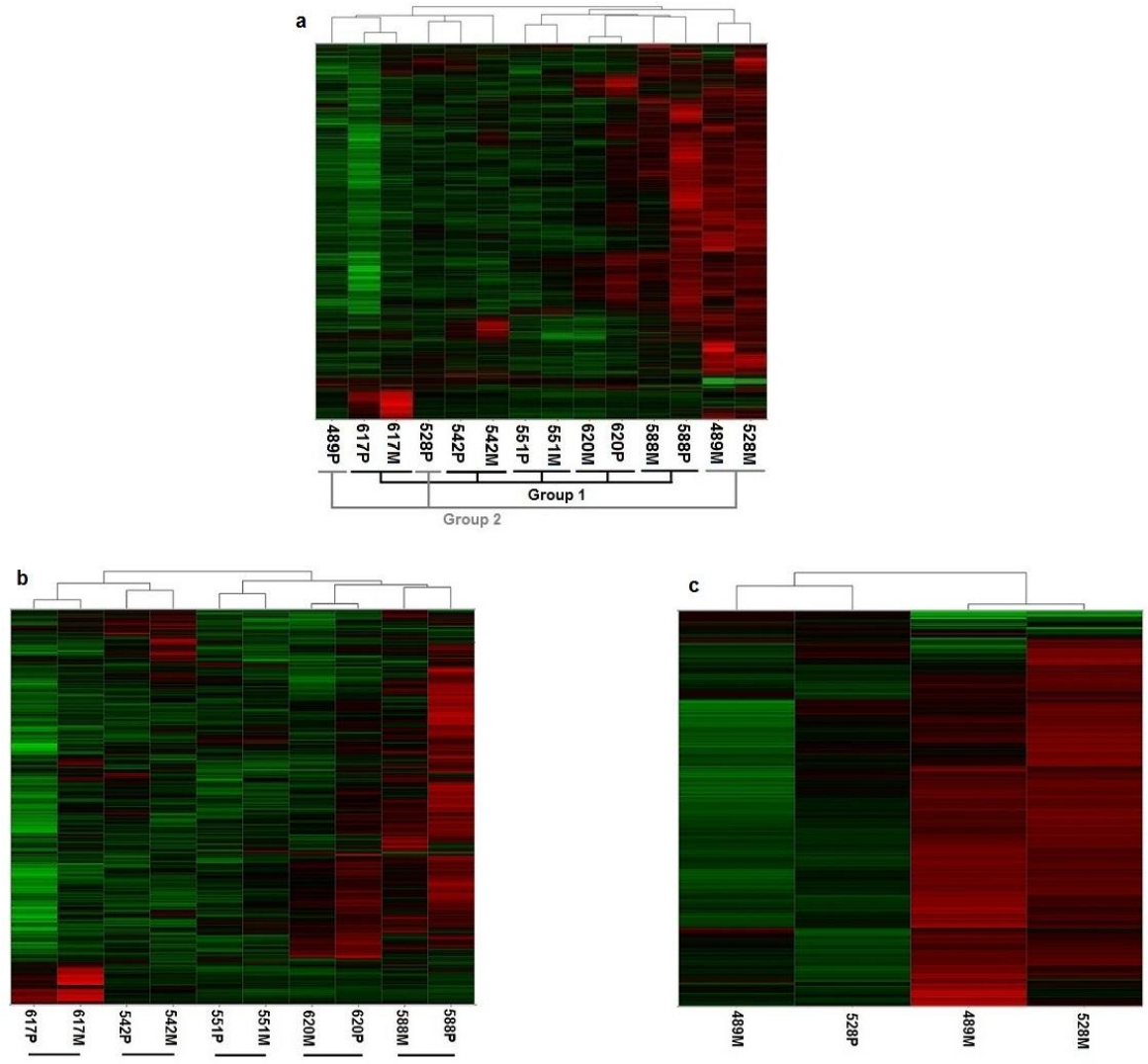


Figure 3.2. Unsupervised classification of differentially expressed genes between primary and metastatic samples identifies two groups of patients. (a) Unsupervised hierarchical clustering performed on 3,365 unique, annotated genes (4,389 probe sets) displaying significant expression variation across all samples (SD $\geq$ 0.8). Primary and metastatic samples from 5 patients (617, 542, 551, 620, 588) clustered closely to one another (Group1) while primary and metastatic samples from 2 patients (489, 528) clustered distantly from one another (Group 2); (b) Unsupervised hierarchical clustering of the same genes/probe sets in (a) across primary (P) and metastatic (M) samples of Group 1 patients. All P samples cluster most closely with their matched M samples for all patients; (c) Unsupervised hierarchical clustering of the same genes/probe sets in (a) across primary (P) and metastatic (M) samples of Group 2 patients. The P samples do not cluster with the matched M samples of the same patient.

Analysis of expression profiles for the 14 primary and metastatic samples identified 3,365 genes (4,389 probe sets) as being significantly differentially expressed (Table A.3.1). Unsupervised hierarchical clustering of these data revealed that the

metastatic samples of five of the seven patients grouped closely with their respective primary samples (Figure 3.2a, b, Group 1). In contrast, the metastatic samples from two of the patients (489M and 528M) clustered most closely with one another and distant from their respective primary samples (Figure 3.2a, c, Group 2). Functional pathway analysis was carried out incorporating the 3,365 differentially expressed genes described above. The results indicated that 13 of the 20 most significantly enriched pathways are associated with EMT or EMT-related functions (*i.e.*, cytoskeleton remodeling, cell adhesion, *etc*., Table 3.1).

Table 3.1. The 20 most significantly enriched pathways between all primary and metastatic samples. Enriched pathways were computed utilizing the 3,365 significantly differentiated expressed genes (4,389 probe sets) represented in the clustering analysis (Figure 3.2). Thirteen of the 20 pathways (highlighted in **bold)** are involved in EMT or EMT-related processes (*i.e.*, cytoskeleton remodeling, cell adhesion, EMT developmental processes, MIF-associated cell adhesion).

	Functional Pathways	**pValue**
1	**Cytoskeleton remodeling_TGF, WNT and cytoskeletal remodeling**	1.53E-13
2	**Cell adhesion_ECM remodeling**	1.03E-11
3	**Cytoskeleton remodeling_Cytoskeleton remodeling**	7.54E-11
4	**Development_Regulation of epithelial-to-mesenchymal transition (EMT)**	9.43E-11
5	**Cell adhesion_Chemokines and adhesion**	1.75E-10
6	Oxidative phosphorylation	1.99E-10
7	**Development_TGF-beta-dependent induction of EMT via MAPK**	2.05E-09
8	DNA damage_Brca1 as a transcription regulator	3.36E-09
9	**Cytoskeleton remodeling_Reverse signaling by ephrin B**	6.64E-09
10	LRRK2 in neurons in Parkinson's disease	2.32E-08
11	Apoptosis and survival_BAD phosphorylation	5.22E-08
12	**Development_TGF-beta-dependent induction of EMT via RhoA, PI3K and ILK**	5.30E-08
13	**Cytoskeleton remodeling_Role of PKA in cytoskeleton reorganisation**	1.30E-07
14	Transcription_Androgen Receptor nuclear signaling	2.02E-07
15	**Immune response_MIF-induced cell adhesion, migration and angiogenesis**	3.06E-07
16	Immune response_MIF - the neuroendocrine-macrophage connector	3.06E-07
17	**Development_TGF-beta-dependent induction of EMT via SMADs**	5.02E-07
18	**Cell adhesion_Integrin-mediated cell adhesion and migration**	6.77E-07
19	Transport_Clathrin-coated vesicle cycle	7.47E-07
20	**Cell adhesion_Role of tetraspanins in the integrin-mediated cell adhesion**	1.27E-06

To explore the possibility that differences in the expression of EMT-associated genes may contribute to the dichotomy between Group 2 primary and metastatic samples, we conducted additional analyses focusing on genes previously established to be directly or indirectly involved in the EMT process (Table A.3.2). Thirty-nine of these EMT-associated genes (61 probe sets) are among the 3,365 genes (4,389 probe sets) used in our clustering analysis (Figure 3.2). Figure 3.3 presents a comparative ranking of these 39 EMT-associated genes with respect to fold-change differences in expression between the Group 1 and Group 2 primary and metastatic samples.

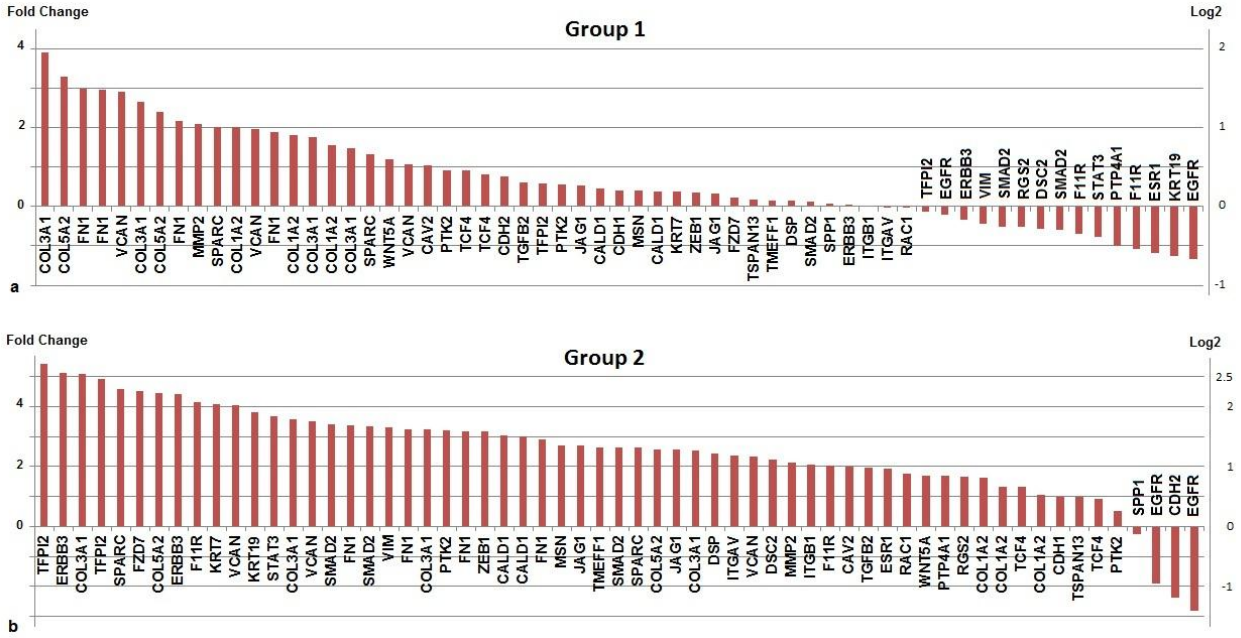


Figure 3.3. Comparative ranking of EMT-associated genes with respect to fold-change differences in expression between the Group 1 and Group 2 primary and metastatic samples. Thirty-nine previously characterized EMT associated genes were identified among the 3,365 genes significantly differentially expressed across 14 tissue samples (see Figure 1). Histograms depict the fold change differences in expression between primary and metastatic samples of Group 1 and Group 2 patients. Although EMT-associated genes are more differentially expressed between primary and metastatic samples from Group 2 than Group 1 patients, Group 1 patients also display large fold differences for some EMT-associated genes. These findings suggest that the observed differences in expression of EMT-associated genes between the primary and metastatic samples of Group 1 vs. Group 2 patients represent differences in a continuum of the EMT-MET process rather than its occurrence in one and absence in the other.

Nearly 74% (45/61) of the EMT-associated probe sets display a > 2-fold change in expression in the Group 2 metastatic samples while only 18% (11/61) display a > 2-fold difference in expression in the Group 1 metastatic samples. In addition, a number of the differences in expression are consistent with the hypothesis that, on the molecular level, Group 2 metastatic samples are more mesenchymal-like than their matched primary samples. For example, the expression of several "mesenchymal biomarkers", *i.e.*, VIM (vimentin), TMEFF1 (transmembrane protein with EGF-like and two follistatin-like domains 1), ITGAV (integrin alpha-V), and ZEB1 (zinc finger E-box-binding homeobox1) are all up-regulated (2 to 3-fold) in Group 2 metastatic samples while being essentially unchanged between Group 1 primary and metastatic samples. However, a number of other genes typically up-regulated during EMT [*e.g.*, COL3A1 (collagen type III alpha-1), FN1 (fibronectin), VCAN (vesican), and MMP2 (matrix metalloproteinase-2] were not only up-regulated in Group 2 metastatic samples but in Group 1 metastatic samples as well, albeit at generally lower levels. These findings suggest that the observed differences in expression of EMT-associated genes between the primary and metastatic samples of Group 1 vs. Group 2 patients represent differences in a continuum of the EMT-MET process rather than its occurrence in one and absence in the other.

To explore this possibility further, we conducted a second clustering analysis using only genes previously implicated in EMT (Table A.3.2). The results presented in Figure 3.4 demonstrate that with respect to these EMT-associated genes, the primary sample of only one patient (620) remained most closely clustered with its metastatic counterpart. The primary and metastatic samples of all of the remaining patients displayed various degrees of isolation from one another.

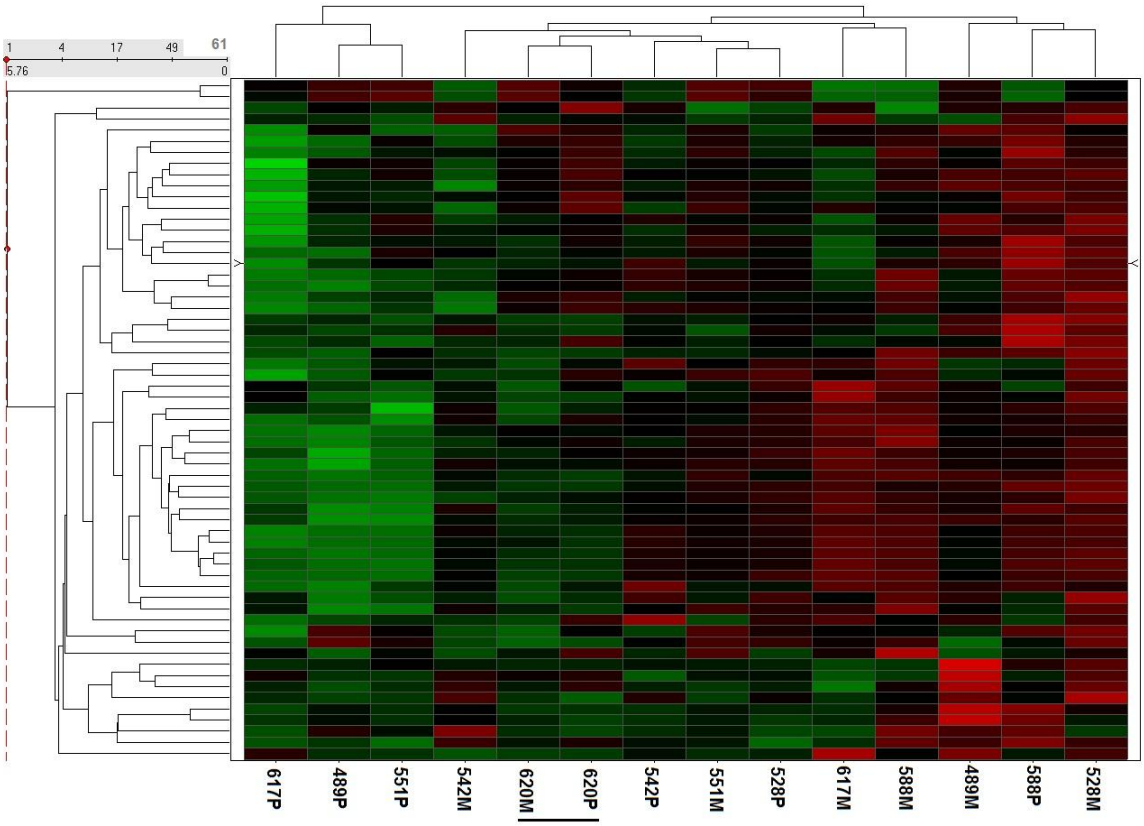


Figure 3.4. Unsupervised classification of 39 differentially expressed EMT associated genes (61 probe sets) demonstrates significant divergence between most primary and metastatic samples. Unsupervised hierarchical clustering of EMT associated genes differentially expressed across all samples demonstrates that the primary and metastatic samples of only one patient (620) are clustered most closely with one another. Primary and metastatic samples of all other patients cluster away from one another consistent with a model whereby all of the metastatic samples have undergone EMT while displaying a range of partial or complete (patient 620) compensating MET transitions at the metastatic site. The alternative hypothesis that metastasis occurs in the absence of EMT is definitively consistent with the molecular profiles of only one Group 1 patient (620).

Collectively, our results are consistent with a model whereby all of the metastatic samples have undergone EMT while displaying a range of partial or complete (patient 620) compensating MET transitions at the metastatic site. The alternative hypothesis that metastasis occurs in the absence of EMT is definitively consistent with the molecular profiles of only one Group 1 patient (620).

The majority of cancer-related deaths are attributable to metastases rather than to primary tumors [1]. For this reason, there is considerable interest in understanding the molecular mechanisms underlying the process with the hope that such knowledge may

lead to more effective therapeutic treatments [12]. Considerable evidence has accumulated in recent years from innumerable *in vitro* and animal model studies e.g., [4-7], indicating that EMT is playing a key role in the metastatic spread of cancer cells from primary sites. Despite the significant body of support for this model, the lack of definitive clinical evidence in human cancer patients has resulted in persistent and spirited skepticism [8, 9]. Much of this skepticism is based upon the fact that human metastatic samples consistently display morphologies indistinguishable from the primary tumors from which they are derived.

In an effort to help resolve this controversy, we conducted morphological and molecular analyses of 14 matched sets of primary and metastatic samples from late staged (III/IV) ovarian cancer patients. Pathological examination revealed no morphological differences between any of the primary and metastatic samples. In contrast, gene expression analyses identified two distinct groups of patient samples. In one group, the molecular profiles of primary and metastatic samples from the same patient displayed indistinguishable molecular profiles. While this result is not inconsistent with an EMT model where mesenchymal-like metastasizing cells have undergone a compensating MET transition at the metastatic site, the results are equally consistent with the hypothesis that metastasis in these patients did not involve EMT. However, molecular profiling also identified a second group of patient samples where the metastatic samples from different patients clustered together and were clearly distinct from their respective primary samples. Further analyses demonstrated that differences in the expression patterns of genes previously associated with EMT clearly separated the primary and metastatic samples isolated from all but one patient. Collectively, our results support a role of EMT in ovarian cancer metastases and demonstrate that indistinguishable morphologies between primary and metastatic cancer samples is not sufficient evidence to negate the role of EMT in the metastatic process.

3.4 References

1. Chaffer, C.L. and R.A. Weinberg, *A perspective on cancer cell metastasis.* Science, 2011. **331**(6024): p. 1559-64.

2. Brabletz, T., *EMT and MET in Metastasis: Where Are the Cancer Stem Cells?* Cancer Cell, 2012. **22**(6): p. 699-701.

3. Valastyan, S. and R.A. Weinberg, *Tumor metastasis: molecular insights and evolving paradigms.* Cell, 2011. **147**(2): p. 275-92.

4. Chaffer, C.L., et al., *Mesenchymal-to-epithelial transition facilitates bladder cancer metastasis: role of fibroblast growth factor receptor-2.* Cancer research, 2006. **66**(23): p. 11271-8.

5. Marian, C.O., et al., *Evidence of epithelial to mesenchymal transition associated with increased tumorigenic potential in an immortalized normal prostate epithelial cell line.* The Prostate, 2011. **71**(6): p. 626-36.

6. Ocana, O.H., et al., *Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer Prrx1.* Cancer Cell, 2012. **22**(6): p. 709-24.

7. Tsai, J.H., et al., *Spatiotemporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis.* Cancer Cell, 2012. **22**(6): p. 725-36.

8. Ledford, H., *Cancer theory faces doubts.* Nature, 2011. **472**(7343): p. 273.

9. Tarin, D., E.W. Thompson, and D.F. Newgreen, *The fallacy of epithelial mesenchymal transition in neoplasia.* Cancer research, 2005. **65**(14): p. 5996-6000; discussion 6000-1.

10. Levayer, R. and T. Lecuit, *Breaking down EMT.* Nature cell biology, 2008. **10**(7): p. 757-9.

11. Thiery, J.P., *Epithelial-mesenchymal transitions in tumour progression.* Nature reviews. Cancer, 2002. **2**(6): p. 442-54.

12. Monteiro, J. and R. Fodde, *Cancer stemness and metastasis: therapeutic consequences and perspectives.* European journal of cancer, 2010. **46**(7): p. 1198-203.

13. Siegel, R., D. Naishadham, and A. Jemal, *Cancer statistics, 2013.* CA: a cancer journal for clinicians, 2013. **63**(1): p. 11-30.

14. Bowen, N.J., et al., *Gene expression profiling supports the hypothesis that human ovarian surface epithelia are multipotent and capable of serving as ovarian cancer initiating cells.* BMC medical genomics, 2009. **2**: p. 71.

VITA

LOUKIA N. LILI

EDUCATION

- **PhD, Biology** (minor: Computational Biology) 2013
  Georgia Institute of Technology, USA
  Thesis Title: "Computational analysis of gene expression profiles of ovarian and pancreatic cancer"
  Supervisor: Professor John F. McDonald
- **Master of Philosophy (MPhil), Mathematics/Mathematical Biology** 2008
  University of Bath, UK
  Thesis Title: "Models in Bacterial Evolution and Bacterial Vector-Borne Diseases"
  Supervisor: Professor Nicholas F. Britton
- **Bachelor of Sciences (BSc), Mathematics** 2005
  University of Crete, Greece
  Undergraduate Thesis Title: "Mathematical Modeling of the Kinetics of Thyroid Hormones"
  Supervisor: Professor Konstadia Lika

PUBLICATIONS

Published

- **Loukia N. Lili**, Lilya V. Matyunina, L. DeEtte Walker, Benedict B. Benigno and John F. McDonald, *Molecular profiling provides evidence for functionally distinct classes of ovarian cancer stroma*, BioMed Research International, 2013. 2013: p.1-9
- **Loukia N. Lili**, N.F.Britton and Edward J.Feil, *The persistence of parasitic plasmids*, Genetics, 2007. 177: p. 399-405

In preparation

- **Loukia N. Lili**, Lilya V. Matyunina, L. DeEtte Walker, George W. Daneker and John F. McDonald, *Evidence for the importance of personalized molecular profiling in pancreatic cancer*
- **Loukia N. Lili,** Lilya V. Matyunina, L. DeEtte Walker, Benedict B. Benigno and John F. McDonald, *Molecular profiling supports the role of EMT in ovarian cancer metastasis*
- Deepraj Ghosh, **Loukia N. Lili**, Daniel J McGrail, Lilya V. Matyunina, John F. McDonald, and Michelle R. Dawson , *Integral Role of PDGF in Mediating TGF-β1-Dependent Mesenchymal Stem Cell Stiffening*

CURRENT RESEARCH INTERESTS

- Analyze and interpret data from microarray and next generation sequencing experiments

- Develop hypotheses on research questions addressed to cancer biology and therapeutics, suggest experimental approaches and analyze and interpret related experimental data

RESEARCH EXPERIENCE

- **Department of Biology, Georgia Institute of Technology, USA** 2009-2013
  Supervisor: Professor John McDonald
  - Developed hypotheses and discovery-based experiments in order to provide a broader understanding of the molecular driven mechanisms in cancer biology and therapeutics. Analyzed and interpreted six independent microarray data sets:
    Three data sets on human genome (two data sets from ovarian primary, metastatic and normal tissue; one data set from pancreatic matched normal and primary cancer tissue);
    One data set from ovarian cell line HEY (transfected with miRNA 429 and control miRNA);
    Two datasets on mouse genome (one from bone-marrow derived mouse cells and another from mouse endothelial cells)
  - Reviewed two manuscripts for publication (PloS ONE; Biology Direct)
- **Department of Mathematics and Department of Biology, University of Bath, UK** 2005-2008
  Supervisors: Professor Nicholas F. Britton, Professor Edward J Feil, and Professor Klaus Kurtenbach
  - Developed and analyzed discrete theoretical mathematical models on bacterial evolution and epidemiology of vector-borne diseases
  - Collaborated with the Lyme disease lab at Bath University, UK and the Lyme disease lab at School of Public Health, Yale. Collaboration involved active participation at weekly lab meetings with oral presentations, as well as field research work of *Ixodes* tick collections at Bath, UK and Yale, CN
- **Department of Mathematics and Department of Biology, University of Crete, Greece** 2003-2004
  Supervisor: Professor Konstadia Lika
  - Developed and analyzed a discrete mathematical model on the kinetics of human thyroidal hormones for Undergraduate Research Thesis

TEACHING EXPERIENCE

- **Department of Biology, Georgia Institute of Technology, USA** 2008-2009
  Prepared, lectured and supervised weekly laboratory classes, graded weekly assignments and final exams for the BIOL1510/1511 course (Biological Principles)
- **Department of Mathematics, Bath University, UK** 2005-2008
  Prepared and lectured weekly classes for assisting undergraduate students with mathematical assignments for various courses. Graded weekly assignments and supervised final exams

COMPUTER SKILLS

- Proficient on Windows and Linux operating systems

- Proficient on R computer language, Bioconductor software for Bioinformatics
- Expert on Expression Console and Spotfire software programs for microarray data analysis

ACADEMIC AWARDS AND HONORS

- **Department of Biology, Georgia Institute of Technology, USA** 2009-2013
  Graduate Research Assistantship
- **Department of Biology, Georgia Institute of Technology, USA** 2008-2009
  Graduate Teaching Assistantship
- **Department of Mathematics, University of Bath, UK** 2005-2008
  Engineering and Physical Sciences Research Council (EPSRC) studentship
- **Imperial College, London** 2007
  Travel Grant for the "Annual Postgraduate Workshop on Ecology"
- **University of Crete, Greece** 2006
  Travel Grant for the "Current Trends in Mathematics Workshop"
- **Department of Mathematics, University of Crete, Greece** 2004
  Graduation Distinction: 3rd Graduated (highest GPA) Student of Class '04
- **Distinctions in Greek Mathematical Olympiads, Greek Mathematical Society** 1997, 1998

CONFERENCES AND WORKSHOPS

- **"**Annual Symposium on Prostate Cancer**",** Clark Atlanta University, Atlanta, GA - **Poster** 2013
- **"**Annual Symposium on Prostate Cancer**",** Clark Atlanta University, Atlanta, GA - Attendant 2011
- "Joint MRC-AACR Conference: Metastasis and the Tumor Microenvironment", Philadelphia, PA - **Poster** 2010
- "European Conference of Mathematical and Theoretical Biology", Edinburgh, Scotland - **Oral Presentation** 2008
- Health Protection Agency, London, UK - **Invited Speaker** 2008
- "Annual meeting of The Society For Mathematical Biology", San Jose, CA - **Poster** 2007
- "Mathematical Models and Experimental Microbial Systems", University of Bath, UK - **Poster** 2007
- "Annual Meeting of The Society For Mathematical Biology", Raleigh, NC - **Poster** 2006
- "Annual Postgraduate Workshop on Ecology", Imperial College, UK - **Oral Presentation** 2006
- "Current Trends In Mathematics", Anogia Academic Village, Crete, Greece - Attendant 2006
- "Mathematics Workshop", University of Crete, Greece - Attendant 2004

LANGUAGES

Fluent in: English, Greek and Spanish

OTHER AWARDS AND HONORS

SPORTS

COMPETITIVE TRIATHLON

- 1st USA Triathlon Southeast Regional Championships (Sprint distance, age group) 2013
- 4th USA Triathlon National Collegiate Championships (Olympic distance, open division) 2012

SAILING

- Captain's Sailing Diploma, Greece 2005
- Level I and Level II Sailing Dinghy Certificates, UK 2006

COMPETITIVE CYCLING

- 7th 500m Greek National Track Championships 2004
- 6th overall, Greek National Road Cup 2003
- 7th MTB Greek National Championships 2003
- 12th overall, International Tour of Rhodes 2003

MUSIC

- Piano Intermediate Degree, Madrigalios Music School, Korinthos, Greece 1997